



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 122907**

**TO: Samuel W Liu**  
**Location: 3c84 / 3c70**  
**Tuesday, May 25, 2004**  
**Art Unit: 1653**  
**Phone: 272-0949**  
**Serial Number: 09 / 880503**

**From: Jan Delaval**  
**Location: Biotech-Chem Library**  
**Rem 1A51**  
**Phone: 272-2504**

**jan.delaval@uspto.gov**

### **Search Notes**

122907

Delaval, Jan

---

**From:** Liu, Samuel  
**Sent:** Tuesday, May 25, 2004 12:43 PM  
**To:** Delaval, Jan  
**Subject:** 09880503

Hi, Jan,

Please search amino acid sequences of SEQ ID NOs: 1, 4, 5, 6, 8 and 9 against commercial protein databases for application 09880503 at least 50 printed results.

Very best,

Samue Liu  
AU 1653, REM 3C84  
571-272-0949



## SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: \_\_\_\_\_ Examiner #: \_\_\_\_\_ Date: \_\_\_\_\_  
Art Unit: \_\_\_\_\_ Phone Number 30 \_\_\_\_\_ Serial Number: \_\_\_\_\_  
Mail Box and Bldg Room Location: \_\_\_\_\_ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

\*\*\*\*\*  
Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc., if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: \_\_\_\_\_

Inventors (please provide full names): \_\_\_\_\_

Earliest Priority Filing Date: \_\_\_\_\_

*\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

\*\*\*\*\*  
STAFF USE ONLY

	Type of Search	Vendors and cost where applicable
Searcher: <u>Jan</u>	NA Sequence (#) _____	STN _____
Searcher Phone #: <u>2284</u>	AA Sequence (#) <input checked="" type="checkbox"/> _____	Dialog _____
Searcher Location _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: <u>5/25</u>	Bibliographic _____	Dr. Link _____
Date Completed: <u>5/25</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep. & Review Time _____	Fulltext _____	Sequence Systems <input checked="" type="checkbox"/> _____
Clerical Prep. Time: <u>20</u>	Patent Family _____	WWW/Internet _____
Online Time: <u>+15</u>	Other _____	Other (specify) _____



C:Species: Homo sapiens (man)  
 C:Date: 21-Sep-1993 #sequence\_revision 25-Aug-1995 #text\_change 08-Dec-2000  
 C:Accession: A46688  
 R: Miyazawa, K.; Shimomura, T.; Kitamura, A.; Kondo, J.; Morimoto, Y.; Kitamura, N.  
 J. Biol. Chem. 268, 10024-10028, 1993  
 A:Title: Molecular cloning and sequence analysis of the cDNA for a human serine protease  
 d coagulation factor XII.  
 A:Reference number: A46688; MUID:93252878; PMID:7683665  
 A:Accession: A46688  
 A:Molecule type: mRNA  
 A:Residues: 1-655 <MIV>  
 A:Cross-references: DBJ:U14012; NID:G219680; PIDN:BA03113.1; PID:9219681  
 A:Experimental source: Liver (mRNA); serum (protein)  
 A:Note: sequence extracted from NCBI backbone (NCBIN:131227, NCBI:131228)  
 A:Note: parts of the sequence, including the amino ends of the heavy and light chains, c  
 C:Genetics:  
 A:Gene: GDB:HGFAC; HGFA; HGFAP  
 A:Cross-references: GDB:9954514  
 A:Map position: 4p16-4p16  
 C:Function:  
 A:Description: activates hepatocyte growth factor by specific proteolytic cleavage  
 A:Pathway: tissue repair and regeneration  
 C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;  
 C:Keywords: glycoprotein; hydrolase; kringle; liver; plasma; serine proteinase  
 F:1-34/Domain: signal sequence #status predicted <SIG>  
 F:108-148/Domain: fibronectin type II repeat homology <1F2>  
 F:164-197/Domain: EGF homology <EG1>  
 F:202-237/Domain: fibronectin type I repeat homology <1F1>  
 F:245-278/Domain: EGF homology <EG2>  
 F:286-367/Domain: kringle homology <KRG>  
 F:373-407/Product: hepatocyte growth factor activator light chain #status experimental  
 F:408-655/Product: hepatocyte growth factor activator heavy chain #status experimental  
 F:408-641/Domain: trypsin homology <TRY>  
 F:40-48-290-468-492-546/Binding site: Carbohydrate (Asn) (covalent) #status predicted  
 F:164-175-169-186-188-197-202-230-228-237-245-256-250-267-269-278-286-367-307-349-338-36  
 F:447-497-538/Active site: His, Asp, Ser #status predicted

Query Match 37.8%; Score 193; DB 1; Length 655;  
 Best Local Similarity 54.5%; Pred. No. 2,9e-13;  
 Matches 36; Conservative 4; Mismatches 26; Indels 0; Gaps 0;  
 QY 3 CYEGNGHYRGKASTDTWGRPCLPWNSATVLTQOYTHAHRSD-ALQGLGKHNCRPNRR 62  
 DB 286 CFLNGTGYRGVASTSGSLCLAWNSDLXQLHLVDSVGAALLGLGPHAYCRNPND 345  
 QY 63 RPWCYV 68  
 DB 346 RPWCYV 351  
 QY 63 RPWCYV 68  
 DB 346 RPWCYV 351

RESULT 20  
 S28941  
 coagulation factor XIIfa (EC 3.4.21.38) - guinea pig (fragment)  
 N:Alternate names: Hageman factor  
 C:Species: Cavia porcellus (guinea pig)  
 C:Date: 25-Feb-1994 #sequence\_revision 03-Aug-1995 #text\_change 21-Jan-2000  
 C:Accession: S28941  
 R:Senba, U.; Yamamoto, T.; Kuniyada, T.; Shibuya, Y.; Tanase, S.; Kambara, T.; Okabe, H.  
 Biochim. Biophys. Acta 1159, 113-121, 1992  
 A:Title: Primary structure of guinea-pig Hageman factor: sequence around the cleavage si  
 A:Reference number: S28941; MUID:93003367; PMID:1390917  
 A:Accession: S28941  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-603 <SEN>  
 A:Cross-references: EMBL:X68615; NID:G49578; PIDN:CAA48600.1; PID:G49579  
 C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;  
 C:Keywords: hydrolase; serine proteinase  
 F:46-87/Domain: fibronectin type II repeat homology <1F2>  
 F:134-165/Domain: fibronectin type I repeat homology <1F1>  
 F:177-208/Domain: EGF homology <EGF>  
 F:216-294/Domain: kringle homology <KRG>  
 F:359-597/Domain: trypsin homology <TRY>

Query Match 35.4%; Score 180.5; DB 2; Length 603;  
 Best Local Similarity 42.9%; Pred. No. 6.2e-12;  
 Matches 36; Conservative 11; Mismatches 32; Indels 5; Gaps 2  
 QY 2 TCYEGNGHYRGKASTDTWGRPCLPWNSATVLTQOYTHAHRSD-ALQGLGKHNCRPN 60  
 DB 215 SCYEGRGVSRYGMARTTVSGAKQRWAS----EATYRNMTAEQALRRGLGHHTFCRNP 270  
 QY 61 RRRPWCYVQGLKPLVQECNVHDC 84  
 DB 271 DTRPWCYVWNGNLSWEYCDLAQC 294

RESULT 21  
 KFHU12  
 coagulation factor XIIfa (EC 3.4.21.38) precursor [validated] - human  
 N:Alternate names: Hageman factor (activated)  
 C:Species: Homo sapiens (man)  
 C:Date: 27-Nov-1985 #sequence\_revision 30-Jun-1991 #text\_change 08-Dec-2000  
 C:Accession: A29411; A26814; A00930; A25191; A22248; A21037  
 R:Cool, D.E.; MacGillivray, R.T.A.  
 J. Biol. Chem. 262, 13662-13673, 1987  
 A:Title: Characterization of the human blood coagulation factor XII gene. Intr  
 A:Reference number: A29411; MUID:88007593; PMID:2888762  
 A:Accession: A29411  
 A:Molecule type: DNA  
 A:Residues: 1-615 <COO>  
 A:Cross-references: GB:M17466; GB:J02807; NID:G180355; PIDN:AA859490.1; PID:G1  
 R:Trippoli, M.; Citarella, F.; Guida, S.; Galeffi, P.; Fantoni, A.; Cortese, R.  
 Nucleic Acids Res. 14, 3146, 1986  
 A:Title: cDNA sequence coding for human coagulation factor XII (Hageman).  
 A:Reference number: A26814; MUID:86176794; PMID:3754331  
 A:Accession: A26814  
 A:Molecule type: mRNA  
 A:Residues: 4-615 <TRI>  
 A:Cross-references: GB:M31315; NID:G182291; PIDN:AAA70225.1; PID:G182292  
 R:Cool, D.E.; Edgell, C.J.S.; Louie, G.V.; Zoller, M.J.; Brayer, G.D.; MacGill  
 J. Biol. Chem. 260, 13666-13676, 1985  
 A:Title: Characterization of human blood coagulation factor XII cDNA. Predicti  
 A:Reference number: A00930; MUID:86033830; PMID:3877053  
 A:Accession: A00930  
 A:Molecule type: mRNA  
 A:Residues: 14-332, 'S', '334-615 <CO2>  
 A:Cross-references: GB:M11723; NID:G180358; PIDN:AAA51986.1; PID:G180359  
 R:Que, B.G.; Davies, E.W., 1986  
 Biochemistry 25, 1525-1528, 1986  
 A:Title: Characterization of a cDNA coding for human factor XII (Hageman facto  
 A:Reference number: A25191; MUID:86216049; PMID:3011063  
 A:Accession: A25191  
 A:Molecule type: mRNA  
 A:Residues: 146-378, 'G', '380-615 <QUE>  
 A:Cross-references: GB:M13147; NID:G180360; PIDN:AAA70224.1; PID:G180361  
 R:McMullen, B.A.; Fujikawa, K.  
 J. Biol. Chem. 260, 5328-5341, 1985  
 A:Title: Amino acid sequence of the heavy chain of human alpha-factor XIIfa (ac  
 A:Reference number: A22248; MUID:85182674; PMID:3886654  
 A:Accession: A22248  
 A:Molecule type: protein  
 A:Residues: 20-379 <MCMA>  
 R:Fujikawa, K.; McMullen, B.A.  
 J. Biol. Chem. 258, 10924-10933, 1983  
 A:Title: Amino acid sequence of human beta-factor XIIfa.  
 A:Reference number: A21037; MUID:83291041; PMID:6604055  
 A:Accession: A21037  
 A:Molecule type: protein  
 A:Residues: 354-362,373-615 <FUJ>  
 R:Harris, R.J.; Ling, V.T.; Spellman, M.W.  
 J. Biol. Chem. 267, 5102-5107, 1992  
 A:Title: O-linked fucose is present in the first epidermal growth factor domain  
 A:Reference number: A44606; MUID:92184750; PMID:1544894  
 A:Contents: annotation; carbohydrate binding site  
 C:Genetics:

```

F:225-226/Cleavage site: His-Ser (plasmin) #status predicted
F:272,321,428/Active site: His, Asp, Ser #status predicted
Query Match 41.8%; Score 213; DB 2; Length 477;
Best Local Similarity 46.4%; Pred. No. 1.4e-15;
Matches 39; Conservative 10; Mismatches 35; Indels 0; Gaps 0;

QY 2 TCYEGNGHFRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRPNDR 61
DB 127 TCYEGGQYTRGTWTAESRVEGINWSSLLTRETYNRMPDAPNLGLGNHNYCRPNPGA 186

QY 62 RRPWCYVQVGLKPLVQECMVHDC 85
DB 187 KPWCYIKAGFTSBCSVPCS 210

RESULT 16
JC5878
plasma hyaluronan-binding protein precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 11-Mar-1998 #sequence_revision 11-Mar-1998 #text_change 16-Jul-1999
C:Accession: JC5878
R:Hashimoto, K.; Tobe, T.; Sumiya, J.; Saguchi, K.; Sano, Y.; Nakano, Y.; Choi-Miura, N.
Biol. Pharm. Bull. 20, 1127-1130, 1997
A:Title: Cloning of the cDNA for a mouse homologue of human PHBP: A novel hyaluronan-bin
A:Reference number: JC5878; MUID:98065239; PMID:9401717
A:Accession: JC5878
A:Molecule type: mRNA
A:Residues: 1-558 <HAS>
C:Comment: This protein acts as serine protease.
C:Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology; trypsin
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-311/Product: plasma hyaluronan-binding protein large chain #status predicted <MATL>
F:75-106/Domain: EGF homology <EG1>
F:113-145/Domain: EGF homology <EG2>
F:152-185/Domain: EGF homology <EG3>
F:192-274/Domain: kringle homology <KRI>
F:312-558/Product: plasma hyaluronan-binding protein small chain #status predicted <MATS>
F:312-558/Domain: trypsin homology <TRY>

Query Match 41.1%; Score 209.5; DB 2; Length 558;
Best Local Similarity 45.8%; Pred. No. 3.9e-15;
Matches 38; Conservative 13; Mismatches 31; Indels 1; Gaps 1;

QY 3 CYEGNGHFRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRPNDR 62
DB 192 CYVGDGYSYRGKSVKVTNQNPCLYNWNSHLLQETYNMFMDAETHGIAHNFCRPNPDGH 251

QY 63 RPPWCYVQVGLKPLVQECMVHDC 84
DB 252 KPWCYVKNSEKWKVEYCDVTVC 274

RESULT 17
A35005
u-plasminogen activator (EC 3.4.21.73) precursor - chicken
N:Alternate names: uPA
C:Species: Gallus gallus (chicken)
C:Date: 20-Jul-1990 #sequence_revision 20-Jul-1990 #text_change 16-Jul-1999
C:Accession: A35005
R:Leslie, N.D.; Kessler, C.A.; Bell, S.M.; Degen, J.L.
J. Biol. Chem. 265, 1339-1344, 1990
A:Title: The chicken urokinase-type plasminogen activator gene.
A:Reference number: A35005; MUID:90110185; PMID:2295632
A:Accession: A35005
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-434 <LES>
A:Cross-references: GB:J05187; NID:g212858; PIDN:AAA49131.1; PID:g212859
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-171/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>

```

```

F:40-71/Domain: EGF homology <EGF>
F:79-158/Domain: kringle homology <KRG>
F:173-428/Product: urokinase-type plasminogen activator chain B #status predicted
F:173-416/Domain: trypsin homology <TRY>
F:162-296,202-218,310-379,342-358,369-397/Disulfide bonds: #status predicted
F:217,272,373/Active site: His, Asp, Ser #status predicted

Query Match 38.9%; Score 198.5; DB 1; Length 434;
Best Local Similarity 54.4%; Pred. No. 4.9e-14;
Matches 37; Conservative 7; Mismatches 15; Indels 5; Gaps 2;

QY 3 CYEGNGHFRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRPNDR 61
DB 79 CYSGNGEDYRGNAEDP----GCLYWDHPSVIRMGDYHADLKNALQLGLGKHNYCRPNRGR 134

QY 62 RRPWCYVQ 69
DB 135 SRPWCYTK 142

RESULT 18
JC4795
plasma hyaluronan-binding protein precursor - human
N:Alternate names: hepatocyte growth factor activator-like protein; PHBP
N:Contains: serine proteinase (EC 3.4.21.-)
C:Species: Homo sapiens (man)
C:Date: 15-Oct-1995 #sequence_revision 16-Aug-1996 #text_change 19-Jul-2002
C:Accession: JC4795
R:Choi-Miura, N.H.; Tobe, T.; Sumiya, J.; Nakano, Y.; Sano, Y.; Mazda, T.; Tom
J. Biochem. 119, 1157-1165, 1996
A:Title: Purification and characterization of a novel hyaluronan-binding prote
A:Reference number: JC4795; MUID:96425001; PMID:8827452
A:Accession: JC4795
A:Molecule type: mRNA
A:Residues: 1-560 <CHO>
A:Cross-references: GB:S83182; NID:G1836158; PIDN:AAB46909.1; PID:gl836159
A:Experimental source: plasma
A:Note: parts of this sequence, including the amino ends of the mature chains,
C:Genetics:
A:Gene: GDB:HABP2; HABP; PHBP; HGFAL
A:Cross-references: GDB:4573962
C:Complex: a disulfide-bonded heterodimer of chains produced from the same pre
C:Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homolo
C:Keywords: chondroitin sulfate proteoglycan; glycoprotein; hyaluronic acid; h
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-313/Product: plasma hyaluronan-binding protein, 50K chain #status predict
F:77-108/Domain: EGF homology <EG1>
F:115-147/Domain: EGF homology <EG2>
F:154-187/Domain: EGF homology <EG3>
F:194-276/Domain: kringle homology <KRI>
F:314-550/Domain: trypsin homology <TRY>
F:314-516/Product: plasma hyaluronan-binding protein, catalytic chain #status
F:54,207/Binding site: carboxylate (Asn) (covalent) #status predicted
F:77-88,82-97,99-108,115-125,120-136,138-147,154-165,159-176,178-187,194-276;
F:362,405,509/Active site: His, Asp, Ser #status predicted

Query Match 37.9%; Score 193.5; DB 1; Length 560;
Best Local Similarity 41.7%; Pred. No. 2.2e-13;
Matches 35; Conservative 15; Mismatches 33; Indels 1; Gaps 2;

QY 3 CYEGNGHFRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRPNDR 62
DB 194 CYVGDGYSYRGKARTVNRQVHACLYWNSHLLQENYNNFMEDAEETHGIGEHNFCEPNDADE 253

QY 63 RPPWCYVQVGLKPLVQECMVHDC 85
DB 254 KPWCYIKVTNDKWKVEYCDVSACS 277

RESULT 19
A45688
hepatocyte growth factor activator (EC 3.4.21.-) precursor [validated] - huma

```

```
F;124-205/Domain: kringle homology <KR1>
F;213-294/Domain: kringle homology <KR2>
F;309-559/Product: t-plasminogen activator chain B #status predicted <BCH>
F;309-553/Domain: trypsin homology <TRY>
F;38-68,66-75,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-289,297-4
F;149,481/Binding site: carbohydrtate (Asn) (covalent) #status predicted
F;308-309/Cleavage site: Arg-ile (plasmin, trypsin) #status predicted
F;355,404,510/Active site: His, Asp, Ser #status predicted

Query Match          44.1%; Score 225; DB 1; Length 559;
Best Local Similarity 47.7%; Pred. No. 7.9e-17;
Matches 41; Conservative 10; Mismatches 35; Indels 0; Gaps 0;

QY      2 TCVEGNGHFYRGKASTDTMGRCPLPWN SATVLQQTTHAHRSDALQLGLGHNYCRNPDR 61
       |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      123 TCEGGQITTVRGTWSTAENGACINWSSALSQRPSARRPNAILGLGNHNYCRNPD 182
               :|||||
QY      62 RRPWCYYVQGLKPLVQCWMVHCADG 87
       :|||||
Db      183 VKPCYVFVKAGKYTTEFCSTPACPKG 208

RESULT 13
JS0600
t-plasminogen activator (EC 3.4.21.68) gamma precursor - common vampire bat
N;Alternate names: tissue plasminogen activator
C;Species: Desmodus rotundus (common vampire bat)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C;Accession: JS0600
R;Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Don
Gene 105, 229-237, 1991
A;Title: The plasminogen activator family from the salivary gland of the vampire bat Des
A;Reference number: JS0597; MUID:92039036; PMID:1937019
A;Accession: JS0600
A;Molecule type: mRNA
A;Residues: 1-394 <KRA>
A;Cross-references: GB:M63990; NID:G166078; PIDN:AAA1595.1; PID:g166079
A;Note: the authors translated the codon ATC for residue 75 as Thr
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-36/Domain: propeptide #status predicted <PRO>
F;37-394/Product: plasminogen activator gamma #status predicted <PA>
F;45-126/Domain: kringle homology <KRG>
F;143-389/Domain: trypsin homology <TRY>
F;45-126,65-108,97-121,131-262,174-190,182-251,276-351,308-324,341-369/Disulfide bonds:
F;142-143/Cleavage site: His-Ser (plasmin) #status predicted
F;189,238,345/Active site: His, Asp, Ser #status predicted
F;315/Binding site: carbohydrtate (Asn) (covalent) #status predicted

Query Match          43.1%; Score 220; DB 2; Length 394;
Best Local Similarity 45.2%; Pred. No. 2e-16;
Matches 38; Conservative 13; Mismatches 33; Indels 0; Gaps 0;

QY      2 TCVEGNGHFYRGKASTDTMGRCPLPWN SATVLQQTTHAHRSDALQLGLGHNYCRNPDR 61
       |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      44 TCYQGVTTVRGTWTSTSEGAQCINNWSNLIRITNGRMPEAVKLGLGNHNYCRNP 103
               :|||||
QY      62 RRPWCYYVQGLKPLVQCWMVHCDA 85
       :|||||
Db      104 SKPCYVIKARKFTSVCSPVCS 127

RESULT 14
A29941
t-plasminogen activator (EC 3.4.21.68) precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A29941; S48205; S48207; S48206
R;Rickles, R.J.; Darrow, A.L.; Strickland, S.
J. Biol. Chem. 263, 1563-1569, 1988
A;Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mR
A;Reference number: A29941; MUID:88087303; PMID:2826484
```

A:Accession: A29941  
A:Molecule type: mRNA  
A:Residues: 1-559 <RIC>  
A:Cross-references: GB:J03520; NID:g202109; PIDN:AAA40470.1; PID:g202110  
R:Blijnen, H.R.; van Heef, B.; Beelen, V.; Collen, D.  
Eur. J. Biochem. 224, 863-871, 1994  
A:Title: Characterization of the murine plasma fibronolytic system.  
A:Reference number: S48202; MUID:95010076; PMID:7523120  
A:Accession: S48205  
A:Molecule type: protein  
A:Residues: 33-37, 'X', 39-40 <LIU>  
A:Accession: S48207  
A:Molecule type: protein  
A:Residues: 309-316 <LI2>  
A:Accession: S48206  
A:Molecule type: protein  
A:Residues: 33-37, 'X', 39-40 <LIW>  
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat  
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F:1-17/Domain: signal sequence #status predicted <SIG>  
F:18-29/Domain: propeptide #status predicted <PRO>  
F:30-559/Product: t-plasminogen activator #status predicted <MAT>  
F:30-308/Product: t-plasminogen activator chain A #status predicted <ACH>  
F:38-75/Domain: fibronectin type I repeat homology <1F1>  
F:83-116/Domain: EGF homology <EGF>  
F:124-205/Domain: kringle homology <K31>  
F:213-294/Domain: kringle homology <K32>  
F:309-559/Product: t-plasminogen activator chain B #status predicted <BCH>  
F:309-559/Domain: trypsin homology <TRY>  
F:38-68,66-75,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265  
F:149,481/Binding site: carboxydrate (Asn) (covalent) #status predicted  
F:308-309/Cleavage site: Arg-ile (plasmin, trypsin) #status predicted  
P:355,404,510/Active site: His, Asp, Ser #status predicted

Query Match 42.9%; Score 219; DB 1; Length 559;  
Best Local Similarity 46.0%; Pred. No. 3.6e-16;  
Matches 40; Conservative 11; Mismatches 36; Indels 0; Gaps 0

QY 2 TCYEGNGHYRCKASTDTMGRLPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDR 61  
DB 123 TCFEEGITYRTGTWTASGAEINWSSVLSPYHARRPNAIKLGLGNHNYCENPDR 182

QY 62 RRPWCYVQGLKPLVQECVWHDCAQK 88  
DB 183 LRPWCYVFRAGKYTFEFGSTPACPKG 209

RESULT 15  
JS0597  
t-plasminogen activator (SC 3.4.21.68) alpha-1 precursor - common vampire bat  
N:Alternate names: tissue plasminogen activator  
C:Species: Desmodus rotundus (common vampire bat)  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999  
C:Accession: JS0597  
R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alago  
Gene 105, 229-237, 1991  
A:Title: The plasminogen activator family from the salivary gland of the vampi  
A:Reference number: JS0597; MUID:92039036; PMID:1937019  
A:Accession: JS0597  
A:Molecule type: mRNA  
A:Residues: 1-477 <KRA>  
A:Cross-references: GB:M63987; NID:g166070; PIDN:AAA31591.1; PID:g166071  
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I  
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-36/Domain: propeptide #status predicted <PRO>  
F:37-477/Product: plasminogen activator alpha-1 #status predicted <PLA>  
F:42-79/Domain: fibronectin type I repeat homology <1FA>  
F:87-120/Domain: EGF homology <EGF>  
F:128-209/Domain: kringle homology <KRG>  
F:226-471/Domain: trypsin homology <TRY>  
F:42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265  
F:133,398/Binding site: carboxydrate (Asn) (covalent) #status predicted  
F:133,398/Binding site: carboxydrate (Asn) (covalent) #status predicted

A;Residues: 1-562 <SAS>  
A;Cross-references: EMBL:X07393; NID:G37243; PIDN:CRA30302.1; PID:G37244  
A;Experimental source: fetal lung cells  
R;Kagitan, H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Saito, A.; Bando, H.; Okada, K.; Ma  
FEBS Lett. 189, 145-149, 1985  
A;Title: Expression in *Escherichia coli* of finger-domain lacking tissue-type plasminogen  
A;Reference number: A91343; MUID:85285620; PMID:3896853  
A;Accession: A91343  
A;Molecule type: mRNA  
A;Residues: 1-38, 'G', '86-433, 'E', '435-562 <KAG>  
A;Experimental source: Detroit 562 cells; ATCC 138  
R;Edlund, T.; NY, T.; Ranby, M.; Heden, L.O.; Palm, G.; Holmgren, E.; Josephson, S.  
Proc Natl Acad Sci U.S.A. 80, 349-352, 1983  
A;Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen activator  
A;Reference number: A93951; MUID:83169656; PMID:6572897  
A;Accession: A93951  
A;Molecule type: mRNA  
A;Residues: 251-358 <EDL>  
A;Experimental source: melanoma cells  
R;Pohl, G.; Kallstrom, M.; Bergsdorf, N.; Wallen, P.; Jornvall, H.  
Biochemistry 23, 3701-3707, 1984  
A;Title: Tissue plasminogen activator: peptide analyses confirm an indirectly derived amino acid  
differences.  
A;Reference number: A90488; MUID:85000468; PMID:6433976  
A;Contents: annotation; melanoma cells, partial sequence of residues 36-562, active and  
R;Pohl, G.; Kaplan, L.; Einarsson, M.; Wallen, P.; Jornvall, H.  
FEBS Lett. 168, 29-32, 1984  
A;Title: Differences between uterine and melanoma forms of tissue plasminogen activator.  
A;Reference number: A91322; MUID:84158956; PMID:6538514  
A;Accession: A91322  
A;Molecule type: protein  
A;Residues: 33-45;311-320 <POH>  
A;Experimental source: uterus  
R;Note: in the uterus, cleavage of the activation peptide may also occur after 38-Gln  
R;van Zonneveld, A.J.; Veerman, H.; Pannekoek, H.  
J. Biol. Chem. 261, 14214-14218, 1986  
A;Contents: annotation; A37567; MUID:87033611; PMID:3021732  
R;Verheijen, J.H.; Caspers, M.P.M.; Chang, G.T.G.; de Munk, G.A.W.; Pouwels, P.H.; Enger  
EMBO J. 5, 3525-3530, 1986  
A;Title: Involvement of finger domain and kringle 2 domain of tissue-type plasminogen activator  
A;Reference number: A37568; MUID:87161761; PMID:3030730  
A;Contents: annotation; fibrin binding site  
R;Dodd, I.; Nunn, B.; Robinson, J.H.  
Thromb. Haemost. 59, 523-528, 1988  
A;Title: Isolation, identification and pharmacokinetic properties of human tissue-type plasminogen  
activator  
A;Reference number: A60902; MUID:89044681; PMID:3142086  
A;Contents: annotation; novel forms of expressed recombinant t-PA  
R;Harris, T.J.R.; Patel, T.; Marston, F.A.O.; Little, S.; Emtage, J.S.; Opdenakker, G.;  
Mol. Biol. Med. 3, 279-292, 1986  
A;Title: Cloning of cDNA coding for human tissue-type plasminogen activator and its expression  
A;Reference number: A54645; MUID:86284200; PMID:3090401  
A;Accession: A54645  
A;Molecule type: mRNA  
A;Residues: 1-562 <HAR>  
A;Cross-references: GB:M15518; NID:G190031; PIDN:AAA60111.1; PID:G190032  
A;Note: parts of this sequence were confirmed by peptide sequencing  
R;Reddy, V.B.; Garrazone, A.J.; Sasak, H.; Wei, C.  
DNA 6, 461-472, 1987  
A;Title: Expression of human uterine tissue-type plasminogen activator in mouse cells using  
A;Reference number: I60110; MUID:88054470; PMID:2824147  
A;Accession: I60110  
A;Molecule type: mRNA  
A;Status: translated from GB/EMBL/DBJ  
A;Title: Isolation and characterization of the human tissue-type plasminogen activator cDNA  
A;Reference number: I55232; MUID:85289338; PMID:3161893  
A;Accession: I55232  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA

A;Residues: 1-36 <R2>  
A;Cross-references: GB:M11890; NID:G339837; PIDN:AAA61213.1; PID:G339839  
C;Comment: Cleavage by plasmin or trypsin produces two chains held together by a single  
disulfide bond. t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-Val  
bond. t-PA binds chain A of fibrin by kringle 2 and the fibronectin type I repeat.  
A;Genetics:  
A;Gene: GDB:PLAT  
A;Cross-references: GDB:119496; OMIM:173370  
A;Map position: 8p12-8p12  
A;Introns: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2; 297/1; 362/2; 408/1; 4  
A;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat ho  
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine protease  
F;1-23/Domain: signal sequence #status predicted <SIG>  
F;24-32/Domain: propeptide #status predicted <PRO>  
F;33-562/Product: t-plasminogen activator #status experimental <MAT>  
F;33-310/Product: t-plasminogen activator chain A #status experimental <ACH>  
F;41-78/Domain: fibronectin type I repeat homology <IF1>  
F;86-119/Domain: EGF homology <EGF>  
F;127-208/Domain: kringle homology <KR1>  
F;215-296/Domain: kringle homology <KR2>  
F;311-562/Product: t-plasminogen activator chain B #status experimental <BCH>  
F;311-566/Domain: trypsin homology <TRY>  
F;41-71,69-78,96-97,91-108,110-119,127-208,148-190,179-203,215-296,236-278,267-  
F;152,483/Binding site: carboxylate (Asn) (covalent) #status experimental  
F;210/Binding site: carboxylate (Asn) (covalent) (partial) #status experimental  
F;310-311/Cleavage site: Arg-Ile (plasmin, trypsin) #status experimental  
F;357,406/Active site: His, Asp #status predicted  
F;513/Active site: Ser #status experimental  
Query Match 44.3%; Score 226; DB 1; Length 562;  
Best Local Similarity 47.7%; Pred. No. 6.2e-17;  
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;  
QY 2 TCYEGNGHFVRGKASTDTMGRPCLPWNSTVLOQTYHAHRS DALQLGLGKHNYCRNPDNR 61  
DB 126 TCYEDQGISYRGTWSTAESGAECTNNSSALAKFPYSGRFPDAIRLGLGNHNYCRNPDNR 185  
QY 62 RRPWCYQVGLKPLVQCMVHDCADG 87  
DB 186 SKPCYVFKAGYSGSEFCSTPACSEG 211  
RESULT 12  
A35029  
t-plasminogen activator (EC 3.4.21.68) precursor - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C;Accession: A35029; A31597  
R;Feng, P.; Ohlsson, M.; NY, T.  
J. Biol. Chem. 265, 2022-2027, 1990  
A;Title: The structure of the rat tPA-less rat tissue-type plasminogen activator  
A;Reference number: A35029; MUID:90130448; PMID:2105315  
A;Accession: A35029  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-559 <FEN>  
A;Cross-references: GB:M31197; NID:G207429; PIDN:AAA42261.1; PID:G207431; GB:J0 26  
R;NY, T.; Leonardson, G.; Hsueh, A.J.W.  
DNA 7, 671-677, 1988  
A;Title: Cloning and characterization of a cDNA for rat tissue-type plasminogen  
A;Reference number: A31597; MUID:89170114; PMID:3148445  
A;Accession: A31597  
A;Molecule type: mRNA  
A;Residues: 1-379, 'K', 381-559 <NYT>  
A;Cross-references: GB:M23697; NID:G530159; PIDN:AAA41812.1; PID:G530160  
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat ho  
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F;1-17/Domain: signal sequence #status predicted <SIG>  
F;18-29/Domain: propeptide #status predicted <PRO>  
F;30-559/Product: t-plasminogen activator #status predicted <MAT>  
F;30-308/Product: t-plasminogen activator chain A #status predicted <ACH>  
F;88-75/Domain: fibronectin type I repeat homology <IF1>  
F;83-116/Domain: EGF homology <EGF>

Db 127 TCYKQGVYRGTSWSTSGACINWNSLLTRTYNGRRSDAITGLGNHNYCRNPNN 186  
QY 62 RRPWCYVQGLKPLVQECMVHDC 85  
Db 187 SRPWCYVVKASFIIEFCVPVCS 210

RESULT 9  
JS0598  
t-Plasminogen activator (EC 3.4.21.68) alpha-2 precursor - common vampire bat  
N/Alternate names: tissue plasminogen activator  
C/Species: Desmodus rotundus (common vampire bat)  
C/Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999  
C/Accession: JS0598  
R/Kraatzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Don  
Gene 105, 229-237, 1991  
A/Title: The plasminogen activator family from the salivary gland of the vampire bat Des  
A/Reference number: JS0597; MUID:92039036; PMID:1937019  
A/Accession: JS0598  
A/Molecule type: mRNA  
A/Residues: 1-477 <KRA>  
A/Cross-references: GB:M63988; NID:G166074; PIDN:AAA31593.1; PID:G166075  
C/Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom  
C/Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F/1-21/Domain: signal sequence #status predicted <SIG>  
F/22-36/Domain: propeptide #status predicted <PRO>  
F/37-477/Product: plasminogen activator alpha-2 #status predicted <PLA>  
F/42-79/Domain: fibronectin type I repeat homology <IFA>  
F/87-120/Domain: EGF homology <EGF>  
F/128-209/Domain: kringle homology <KRG>  
F/226-471/Domain: trypsin homology <TRY>  
F/42-72, 70-79, 87-98, 92-109, 111-120, 128-209, 149-191, 180-204, 214-345, 257-273, 265-334, 359-4  
F/185, 398/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F/225-226/Cleavage site: His-Ser (plasmin) #status predicted  
F/272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 47.3%; Score 241; DB 2; Length 477;  
Best Local Similarity 50.0%; Pred. No. 1.2e-18;  
Matches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRGKASTDTWGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNCRNPNN 61  
Db 127 TCYKQGVYRGTSWSTSGACINWNSLLTRTYNGRRSDAITGLGNHNYCRNPNN 186  
QY 62 RRPWCYVQGLKPLVQECMVHDC 85  
Db 187 SRPWCYVVKASFIIEFCVPVCS 210

RESULT 10  
I38098  
t-Plasminogen activator precursor, inactive endothelial splice form - human  
N/Alternate names: tissue plasminogen activator  
C/Species: Homo sapiens (man)  
C/Date: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 22-Jun-1999  
C/Accession: I38098; S01678  
R/Siebert, P.D.; Fong, K.  
Nucleic Acids Res. 18, 1086, 1990  
A/Title: Variant tissue type plasminogen activator (PLAT) cDNA obtained from human endo  
A/Reference number: I38098; MUID:90192128; PMID:1969145  
A/Accession: I38098  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-291 <SIE>  
A/Cross-references: EMBL:X13097; NID:G35282; PIDN:CAA31489.1; PID:G35283  
C/Comment: For the main splice form, see PIR:UKHUT. This form probably does not have pro  
C/Genetics:  
A/Gene: GDB:PLAT  
A/Cross-references: GDB:119496; OMIM:173370  
A/Map position: 8p12-8p12  
A/Introns: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2  
C/Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom

C/Keywords: alternative splicing; fibrinolysis; glycoprotein; kringle  
F/1-23/Domain: signal sequence #status predicted <SIG>  
F/24-32/Domain: propeptide #status predicted <PRO>  
F/33-291/Product: t-plasminogen activator, inactive endothelial splice form #s  
F/41-78/Domain: fibronectin type I repeat homology <IFA>  
F/86-119/Domain: EGF homology <EGF>  
F/127-208/Domain: kringle homology <KRI>  
F/215-291/Domain: kringle homology #status atypical <XR2>  
F/41-71,69-78,86-97,91-108,110-119,127-208,148-190,179-203/Disulfide bonds: #s  
us pr

Query Match 44.3%; Score 226; DB 2; Length 291;  
Best Local Similarity 47.7%; Pred. No. 3.3e-17;  
Matches 41; Conservative 8; Mismatches 37; Indels 0; Gaps 0

QY 2 TCYEGNGHYRGKASTDTWGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNCRNPNN 61  
Db 126 TCYEDQGISYRGTSWSTSGAECTNWNSSALQAQVSGRRPDAILGLGNHNYCRNPDD 186  
QY 62 RRPWCYVQGLKPLVQECMVHDC 87  
Db 186 SRPWCYVVKAGYSSBPFCSTPACSEG 211

RESULT 11  
UKHUT  
t-Plasminogen activator (EC 3.4.21.68) precursor [validated] - human  
N/Alternate names: t-PA; tissue plasminogen activator  
C/Species: Homo sapiens (man)  
C/Date: 14-Nov-1983 #sequence\_revision 14-Nov-1983 #text\_change 08-Dec-2000  
C/Accession: A94004; A23529; JT0562; A91323; A91343; A93951; A91322; A  
R/NY, T.; Elgh, F.; Lund, B.  
Proc. Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984  
A/Title: The structure of the human tissue-type plasminogen activator gene: co  
A/Reference number: A94004; MUID:84298137; PMID:6089198  
A/Accession: A94004  
A/Molecule type: DNA  
A/Residues: 1-562 <NYT>  
A/Cross-references: GB:L00141  
A/Note: the codon given for residue 93 (ACC) is inconsistent with the authors' ansle  
R/Friezner Degen, S.J.; Rajput, B.; Reich, E.  
J. Biol. Chem. 261, 6972-6985, 1986  
A/Title: Purification and characterization of tissue plasminogen activator sec  
A/Reference number: A23529; MUID:86196143; PMID:3009482  
A/Accession: A23529  
A/Molecule type: DNA  
A/Residues: 1-562 <DEG>  
A/Cross-references: GB:K03021; NID:G339817; PIDN:AAA98809.1; PID:G339818  
R/Itagaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuda, S.; Higashio, K.  
Agric. Biol. Chem. 55, 1225-1232, 1991  
A/Title: Purification and characterization of tissue plasminogen activator sec  
A/Reference number: JT0562; MUID:91291340; PMID:1368681  
A/Accession: JT0562  
A/Molecule type: mRNA  
A/Residues: 31-562 <ITA>  
A/Cross-references: DBJ:D01096; NID:G220128; PIDN:BA000881.1; PID:G441174  
A/Experimental source: embryonic lung fibroblast IMR-90 cells  
A/Note: part of this sequence, including the amino end of the mature protein, s conf  
R/Pennica, D.; Holmes, W.E.; Kohr, W.J.; Harkins, R.N.; Vehar, G.A.; Ward, C.A  
Nature 301, 214-221, 1983  
A/Title: Cloning and expression of human tissue-type plasminogen activator cDN  
A/Reference number: A93293; MUID:63115262; PMID:6337343  
A/Accession: A93293  
A/Molecule type: mRNA  
A/Residues: 1-562 <PEN>  
A/Cross-references: GB:L00141  
A/Experimental source: melanoma cells  
R/Sasaki, H.; Saito, Y.; Hayashi, M.; Otsuka, K.; Niwa, M.  
Nucleic Acids Res. 15, 5695, 1988  
A/Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA fro  
A/Reference number: S02125; MUID:88262579; PMID:3133640  
A/Accession: S02125  
A/Status: translation not shown  
A/Molecule type: mRNA



C; Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase  
F;1-20/Domain: signal sequence #status predicted <SIG>  
F;21-179/Product: plasminogen activator chain A #status predicted <MA1>  
F;21-179/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>  
F;33-64/Domain: EGF homology <EGF>  
F;72-153/Domain: kringle homology <KR>  
F;181-433/Product: plasminogen activator chain B #status predicted <MA2>  
F;181-421/Domain: trypsin homology <TRY>  
F;170-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted  
F;226,277,378/Active site: His, Asp, Ser #status predicted

Query Match 77.6%; Score 396; DB 1; Length 433;  
Best Local Similarity 76.1%; Pred. No. 1.3e-35;  
Matches 67; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFGKASTDTMGRCPLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60  
Db 70 KTCYQNGHSGYRGKANDLSGRPCPLAWDSPTVLLKMYHAHRSDALQLGLGKHNYCRNPDN 129

QY 61 RRRPWCYVQGLKPLVQECWVHDCADGK 88  
Db 130 QRRPWCYVQGLKPLVQECWVHDCVSGK 157

RESULT 6  
UKMS

u-plasminogen activator (EC 3.4.21.73) precursor - mouse  
C; Species: Mus musculus (house mouse)  
C; Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 18-Jun-1999  
C; Accession: A29420; A24615  
R; Degen, S.J.F.; Heckel, J.L.; Reich, E.; Degen, J.L.  
Biochemistry 26, 8270-8279, 1987  
A; Title: The murine urokinase-type plasminogen activator gene.  
A; Reference number: A29420; MUID:86163489; PMID:2831940  
A; Accession: A29420  
A; Molecule type: DNA  
A; Residues: 1-433 <DEG>  
A; Cross-references: GB:M17922; NID:G202296; PIDN:AAA40539.1; PID:G202297  
R; Balin, D.; Vassalli, J.D.; Combepine, C.; Godeau, F.; Nagamine, Y.; Reich, E.; Kocher, E.; J. Biochem. 148, 225-232, 1985  
A; Title: Cloning, nucleotide sequencing and expression of cDNAs encoding mouse urokinase  
A; Reference number: A24615; MUID:85179474; PMID:2985383  
A; Accession: A24615  
A; Molecule type: mRNA  
A; Residues: 1-433 <BEL>  
A; Cross-references: GB:X02389; NID:G55127; PIDN:CAA26231.1; PID:G55128  
C; Genetics:  
A; Introns: 19/3; 30/1; 66/1; 124/2; 155/1; 229/2; 279/1; 326/1; 375/3  
C; Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try  
C; Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase  
F;1-20/Domain: signal sequence #status predicted <SIG>  
F;21-178/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>  
F;32-63/Domain: EGF homology <EGF>  
F;71-152/Domain: kringle homology <KR>  
F;180-433/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>  
F;180-421/Domain: trypsin homology <TRY>  
F;169-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status predicted  
F;226,277,378/Active site: His, Asp, Ser #status predicted

Query Match 75.9%; Score 387; DB 1; Length 433;  
Best Local Similarity 72.7%; Pred. No. 1.2e-34;  
Matches 64; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFGKASTDTMGRCPLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60  
Db 69 KTCYHNGDSYRGKANDTGRPCPLAWNAPVLOQTYHAHRSDALQLGLGKHNYCRNPDN 128

QY 61 RRRPWCYVQGLKPLVQECWVHDCADGK 88  
Db 129 QRRPWCYVQGLKPLVQECWVHDCVSLK 156

RESULT 7

JS0599  
t-plasminogen activator (EC 3.4.21.68) beta precursor - common vampire bat  
N; Alternate names: tissue plasminogen activator  
C; Species: Desmodus rotundus (common vampire bat)  
C; Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999  
C; Accession: JS0599  
R; Kraetzschmar, J.; Haendler, B.; Langer, G.; Boisdol, W.; Bringmann, P.; Alagon  
Gene 105, 229-237, 1991  
A; Title: The plasminogen activator family from the salivary gland of the vampire  
A; Reference number: JS0597; MUID:92039036; PMID:1937019  
A; Accession: JS0599  
A; Molecule type: mRNA  
A; Residues: 1-431 <KRA>  
A; Cross-references: GB:M63989; NID:G166076; PIDN:AAA31594.1; PID:G166077  
C; Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I  
C; Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;22-36/Domain: propeptide #status predicted <PRO>  
F;37-431/Product: plasminogen activator beta #status predicted <PLA>  
F;41-74/Domain: EGF homology <EGF>  
F;82-163/Domain: kringle homology <KR>  
F;180-425/Domain: trypsin homology <TRY>  
F;41-52,46-63,65-74,82-163,103-145,134-158,168-299,211-227,219-288,313-388/Disu  
F;139,352/Binding site: carboxylate (Asn) (covalent) #status predicted  
F;179-180/Cleavage site: His-Ser (plasmin) #status predicted  
F;226,275,382/Active site: His, Asp, Ser #status predicted  
F;345-361,378-406/Disulfide bonds: #status predicted

Query Match 47.3%; Score 241; DB 2; Length 431;  
Best Local Similarity 50.0%; Pred. No. 1.1e-18;  
Matches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0;

QY 2 TCYEGNGHFGKASTDTMGRCPLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61  
Db 81 TCYKQGVYRGVTSSTESGAQCINNSNLLTRTYNGRSDAITLGLGHNHYCRNPDN 140

QY 62 RRPWCYVQGLKPLVQECWVHDC 85  
Db 141 SKPWCYVTKASKFILEFCSEVPCS 164

RESULT 8  
A34369

t-plasminogen activator (EC 3.4.21.68) precursor - false vampire bat (Megaderma  
C; Species: Megaderma lyra  
C; Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C; Accession: A34369  
R; Gardell, S.J.; Duong, L.T.; Diehl, R.E.; York, J.D.; Hare, T.R.; Register, R.  
J. Biol. Chem. 264, 17947-17952, 1989  
A; Title: Isolation, characterization, and cDNA cloning of a vampire bat salivar  
A; Reference number: A34369; MUID:90036867; PMID:2509450  
A; Accession: A34369  
A; Status: preliminary  
A; Molecule type: mRNA  
A; Residues: 1-477 <GAR>  
A; Cross-references: GB:J05082; NID:G166080; PIDN:AAA31596.1; PID:G166081  
C; Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I  
C; Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;22-36/Domain: propeptide #status predicted <PRO>  
F;37-477/Product: plasminogen activator #status predicted <PLA>  
F;42-79/Domain: fibronectin type I repeat homology <1FA>  
F;87-120/Domain: EGF homology <EGF>  
F;128-209/Domain: kringle homology <KR>  
F;226-471/Domain: trypsin homology <TRY>  
F;42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-  
F;272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 47.3%; Score 241; DB 1; Length 477;  
Best Local Similarity 50.0%; Pred. No. 1.2e-18;  
Matches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0;

QY 2 TCYEGNGHFGKASTDTMGRCPLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61



C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase  
F:1-20/Domain: signal sequence #status predicted <SIG>  
F:21-176/Product: plasminogen activator chain A #status predicted <ACH>  
F:30-61/Domain: EGF homology <EGF>  
F:69-150/Domain: kringle homology <KRG>  
F:178-433/Product: plasminogen activator chain B #status predicted <BCH>  
F:178-433/Domain: trypsin homology <TRY>  
F:167-298,298-224,216-287,315-384,347-363,374-402/Disulfide bonds: #status predicted  
F:223,274,378/Active site: His, Asp, Ser #status predicted  
F:324/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 93.1%; Score 475; DB 1; Length 433;  
Best Local Similarity 94.3%; Pred. No. 2.9e-44;  
Matches 83; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRNPDN 60  
Db 67 KTCYEGNGHFYRGKASTDTMGPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRNPDN 125

QY 61 RRRPWCYQVGLKPLVQECVWHDGK 88

Db 127 RRRPWCYQVGLKPLVQECVWHDGK 154

#### RESULT 3

UPLG

u-plasminogen activator (EC 3.4.21.73) precursor - pig

N:Alternate names: uPA

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 04-Dec-1986 #sequence\_revision 17-Mar-1987 #text\_change 07-Aug-1998

C:Accession: A00932

R:Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, E.

Nucleic Acids Res. 12, 9525-9541, 1984

A:Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.

A:Reference number: A00932; MUID:85087954; PMID:6096832

A:Accession: A00932

A:Molecule type: DNA

A:Residues: 1-240, 'H', 242-442 <NAG1>

A:Experimental source: kidney cell line LLC-PK1

R:Nagamine, Y.

submitted to the Protein Sequence Database, December 1986

A:Reference number: A37566

A:Contents: annotation; correction to residue 241

C:Genetics:

A:Introns: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3

C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try

F:1-20/Domain: signal sequence #status predicted <SIG>

F:21-189/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>

F:33-64/Domain: EGF homology <EGF>

F:72-153/Domain: kringle homology <KRG>

F:190-442/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>

F:152/Binding site: trypsin homology <TRY>

F:179-310,220-236,228-299,324-393,356-372,383-411/Disulfide bonds: #status predicted

F:235,286,387/Active site: His, Asp, Ser #status predicted

Query Match 82.4%; Score 420; DB 1; Length 442;  
Best Local Similarity 81.8%; Pred. No. 3e-38;  
Matches 72; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRNPDN 60  
Db 70 QTCFEGNGHSYRGKANTDTMGPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRNPDN 129

QY 61 RRRPWCYQVGLKPLVQECVWHDGK 88

Db 130 QRRPWCYQVGLKPLVQECVWHDGK 157

#### RESULT 4

S18932

u-plasminogen activator (EC 3.4.21.73) precursor - rat

N:Alternate names: plasminogen activator, urokinase-type; urinary plasminogen  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 18-Oct-1989 #sequence\_revision 10-Feb-1995 #text\_change 18-Jun-1999  
C:Accession: S24604; I60186; I53472; S18932  
R:Rabbani, S.A.

submitted to the EMBL Data Library, April 1992

A:Reference number: S24604

A:Accession: S24604

A:Molecule type: mRNA

A:Residues: 1-15, 'H', 17-23, 'G', 25-331, 'N', 333-432 <RAB>

A:Cross-references: EMBL:X65651; NID:G57456; PIDN:CAA46601.1; PID:G57457

A:Experimental source: tissue kidney

R:Henderson, B.R.; Tansey, W.F.; Phillips, S.M.; Ramshaw, I.A.; Kefford, R.F.

Cancer Res. 52, 2489-2496, 1992

A:Title: Transcriptional and posttranscriptional activation of urokinase plasminogen

A:Reference number: I60186; MUID:92233409; PMID:1568219

A:Accession: I60186

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-432 <RES>

A:Cross-references: EMBL:X63434; NID:G57465; PIDN:CAA45028.1; PID:G57466

A:Experimental source: strain Fischer 344; tissue mammary

R:Ragno, P.; Cassano, S.; Degen, J.; Kessler, C.; Blasi, F.; Rossi, G.

FEBS Lett. 306, 193-198, 1992

A:Title: The receptor for the plasminogen activator of urokinase type is up-regulated

A:Reference number: I53472; MUID:92339549; PMID:1321734

A:Accession: I53472

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 31-62 <RE2>

A:Cross-references: EMBL:X66907; NID:G396200; PIDN:CAA47356.1; PID:G938279

C:Genetics:

A:Gene: uPA

C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t

C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-177/Product: urokinase-type plasminogen activator chain A #status predicted

F:31-62/Domain: EGF homology <EGF>

F:70-151/Domain: kringle homology <KRG>

F:179-432/Product: urokinase-type plasminogen activator chain B #status predicted

F:179-420/Domain: trypsin homology <TRY>

F:168-300,210-226,218-289,314-383,346-362,373-401/Disulfide bonds: #status predicted

F:225,276,377/Active site: His, Asp, Ser #status predicted

Query Match 79.6%; Score 406; DB 1; Length 432;  
Best Local Similarity 79.5%; Pred. No. 1e-36;  
Matches 70; Conservative 5; Mismatches 13; Indels 0; Gaps 0

QY 1 KTCYEGNGHFYRGKASTDTMGPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRNPDN 60  
Db 68 KTCYHGNGSYRGKANTDTKGRPCLANSPAVLQQTTHAHRSDALSGLGKHNYCRNPDN 127

QY 61 RRRPWCYQVGLKPLVQECVWHDGK 88

Db 128 QRRPWCYQVGLKPLVQECVWHDGK 155

#### RESULT 5

JN0560

u-plasminogen activator (EC 3.4.21.73) precursor - bovine

N:Alternate names: uPA

C:Species: Bos primigenius taurus (cattle)

C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 16-Jul-1999

C:Accession: JN0560

R:Kraetzschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schleuning, W.D.

Gene 125, 177-183, 1993

A:Title: Bovine urokinase-type plasminogen activator and its receptor: cloning and inc

A:Reference number: JN0560; MUID:93216119; PMID:8385052

A:Accession: JN0560

A:Molecule type: mRNA

A:Residues: 1-433 <KRA>

A:Cross-references: GB:L03546; NID:G163800; PIDN:AAA51419.1; PID:G163801

C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t

A;Accession: I52209  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 145-161 <NAG1>  
A;Cross-references: GB:K03027; NID:G340174; PIDN:AAA61257.1; PID:G340175  
R;Agar, M.; Hikamatsu, R.; Kaneda, T.; Hayasuke, N.; Arimura, H.; Nishida, M.; Suyama, Gene 36, 183-188, 1985  
A;Title: Molecular cloning of cDNA coding for human preprourokinase.  
A;Accession: J0102; MUID:86056954; PMID:2415429  
A;Molecule type: mRNA  
A;Residues: 1-213, 'I', 215-431 <NAG2>  
A;Cross-references: GB:K03226; NID:G340155; PIDN:AA097138.1; PID:G340158; GB:D00244; NID:G340159  
P;Verde, P.; Stoppelli, M.P.; Galeffi, P.; Di Nocera, P.; Blasi, F.  
Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984  
A;Title: Identification and primary sequence of an unspliced human urokinase poly(A)+ RNA  
A;Reference number: A37561; MUID:84272706; PMID:6589620  
A;Accession: A37561  
A;Molecule type: mRNA  
A;Residues: 66-431 <VER>  
A;Cross-references: GB:D00244; NID:G220138  
R;Jacobs, P.; Cravador, A.; Loriau, R.; Brockly, F.; Colau, B.; Chuchana, P.; van Elsen, DNA 4, 139-146, 1985  
A;Title: Molecular cloning, sequencing, and expression in *Escherichia coli* of human prepro-urokinase.  
A;Reference number: I38102; MUID:85203359; PMID:3888571  
A;Accession: I38102  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-150, 'I', 152-213, 'I', 215-385, 'C', 387-429, 'V', 431 <JAC>  
A;Cross-references: EMBL:X02760; NID:G35297; PIDN:CAA2535.1; PID:G35298  
R;Yoshimoto, M.; Ushiyama, Y.; Sakai, M.; Tamaki, S.; Hara, H.; Takahashi, K.; Sawasaki, Biochim. Biophys. Acta 1293, 83-89, 1996  
A;Title: Characterization of single chain urokinase-type plasminogen activator with a novel amino acid sequence.  
A;Reference number: S65783; MUID:96186279; PMID:8652631  
A;Accession: S65783  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 21-140, 'L', 142-213, 'I', 215-431 <YOS>  
A;Cross-references: EMBL:D11143; NID:G311467; PIDN:BA01919.1; PID:G1199928  
R;Gunzler, W.A.; Steffens, G.J.; Otting, F.; Kim, S.W.A.; Frankus, E.; Flohe, L. Hoppe-Seyler's Z. Physiol. Chem. 363, 1155-1165, 1982  
A;Title: The primary structure of high molecular mass urokinase from human urine.  
A;Reference number: A37562; MUID:83055084; PMID:6754569  
A;Accession: A37562  
A;Molecule type: protein  
A;Residues: 21-177 <GUNS>  
R;Schaller, J.; Nick, H.; Rickli, E.E.; Gillesen, D.; Lergier, W.; Studer, R.O. Eur. J. Biochem. 125, 251-257, 1982  
A;Title: Human low-molecular-weight urinary urokinase. Partial characterization and preliminary sequence.  
A;Reference number: A37563; MUID:83003608; PMID:6749491  
A;Accession: A37563  
A;Molecule type: protein  
A;Residues: 156-176, 179-193, 'T', 195, 'T', 197-224 <SCH>  
R;Steffens, G.J.; Gunzler, W.A.; Otting, F.; Frankus, E.; Flohe, L. Hoppe-Seyler's Z. Physiol. Chem. 363, 1043-1058, 1982  
A;Title: The complete amino acid sequence of low molecular mass urokinase from human urine.  
A;Reference number: A37564; MUID:83055099; PMID:6754572  
A;Accession: A37564  
A;Molecule type: protein  
A;Residues: 158-410 <ST>  
R;Kantzer, E.J.; Buko, A.; Menon, G.; Sarin, V.K. Biochem. Biophys. Res. Commun. 171, 401-406, 1990  
A;Title: Identification of a fucose and attempt to determine its attachment site  
A;Reference number: A35689; MUID:90365737; PMID:2393398  
A;Accession: A35689  
A;Molecule type: protein  
A;Residues: 21-30, 'X', 32, 'X', 34-38, 'X', 40-43 <KEN>  
A;Note: identification of a fucose and attempt to determine its attachment site  
R;Rabbani, S.A.; Desjardins, J.; Bell, A.W.; Banville, D.; Mazar, A.; Henkin, J.; Goltz, Biochem. Biophys. Res. Commun. 173, 1058-1064, 1990  
A;Title: An amino-terminal fragment of urokinase isolated from a prostate cancer cell line.  
A;Reference number: A36697; MUID:91097529; PMID:2125213  
A;Accession: A36697

A;Molecule type: protein  
A;Residues: 21-34 <RAB>  
R;Li, X.; Bokman, A.M.; Illinas, M.; Smith, R.A.G.; Dobson, C.M. submitted to the Brookhaven Protein Data Bank, July 1993  
A;Reference number: A51255; PDB:1KDU and disulfide bond assignments by (1)H-NMR esidu  
A;Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR  
R;Li, X.; Smith, R.A.G.; Dobson, C.M. Biochemistry 31, 9562-9571, 1992  
A;Title: Sequential (1)H NMR assignments and secondary structure of the kringle main  
A;Reference number: A44375; MUID:9303110; PMID:1337118  
R;Hansen, A.P.; Petros, A.M.; Meadows, R.P.; Nettekheim, D.G.; Mazar, A.P.; Ole submitted to the Brookhaven Protein Data Bank, January 1994  
A;Reference number: A66822; PDB:1URK  
A;Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR esidu  
R;Spraggon, G.S.; Phillips, C.; Nowak, U.K.; Fonting, C.P.; Saunders, D.; Dobson C.M. submitted to the Brookhaven Protein Data Bank, July 1995  
A;Reference number: A66058; PDB:1LMW  
A;Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 169-175; 3-426  
C;Comment: This enzyme is found in urine in a high molecular mass form, consist of 7  
C;Comment: Urokinase-type plasminogen activator proteolytically activates plasminogen,  
C;Genetics:  
A;Gene: GDB:PLAU  
A;Cross-references: GDB:119497; OMIM:191840  
A;Map position: 10q24-10q24  
A;Introns: 19/3; 29/1; 65/1; 123/2; 154/1; 227/2; 277/1; 324/1; 373/3  
C;Function:  
A;Description: proteolytically activates plasminogen  
A;Pathway: fibrinolysis  
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology  
C;Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serine  
F;1-20/Domain: signal sequence #status predicted <sig>  
F;21-431/Product: urokinase-type plasminogen activator, single chain form #status predicted  
F;21-177/Product: urokinase-type plasminogen activator chain A #status experimental  
F;31-62/Domain: EGF homology <EGF>  
F;70-151/Domain: kringle homology <KRG>  
F;156-177/Product: urokinase-type plasminogen activator chain A1 #status experimental  
F;179-431/Product: urokinase-type plasminogen activator chain B #status experimental  
F;179-419/Domain: tryptase homology <TRY>  
F;31-39,33-51,53-62,70-151,91-133,122-146,168-239,209-225,217-288,313-382,345-372-  
F;38/Binding site: carboxylate (Thr) (covalent) #status predicted  
F;178-179/Cleavage site: Lys-Ile (glutamine) #status experimental  
F;224,275,376/Active site: His, Asp, Ser #status experimental  
F;322/Binding site: carboxylate (Asn) (covalent) #status experimental  
Query Match 100.0%; Score 510; DB 1; Length 431;  
Best Local Similarity 100.0%; Pred. No. 4,3e-48; Mismatches 0; Gaps 0  
Matches 88; Conservative 0; Indels 0  
QY 1 KTCYEGNGHFYRGKASTDNGRPLPNSATVLOQTYHAHRSALQGLGKHNYCRNPDN 60  
Db 68 KTCYEGNGHFYRGKASTDNGRPLPNSATVLOQTYHAHRSALQGLGKHNYCRNPDN 127  
QY 61 RRRPCYVQVGLKPLVQECMVHDCADGK 88  
Db 128 RRRPCYVQVGLKPLVQECMVHDCADGK 155  
RESULT 2  
UKBAY  
u-plasminogen activator (BC 3.4.21.73) precursor - yellow baboon  
C;Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)  
C;Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 18-Jun-1999  
C;Accession: S14687; MUID:90287734; PMID:2113276  
R;Au, Y.P.T.; Wang, T.W.; Clowes, A.W. Nucleic Acids Res. 18, 3411, 1990  
A;Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type  
A;Reference number: S14687; MUID:90287734; PMID:2113276  
A;Accession: S14687  
A;Molecule type: mRNA  
A;Residues: 1-433 <NUY>  
A;Cross-references: EMBL:X51935; NID:G38130; PIDN:CAA36200.1; PID:G38131  
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology;



R;Malgaretti, N.; Acquati, P.; Magnaghi, P.; Bruno, L.; Pontoglio, M.; Rocchi, M.; Sacco  
Proc. Natl. Acad. Sci. U.S.A. 89, 11584-11588, 1992  
A;Title: Characterization by yeast artificial chromosome cloning of the linked apolipoprotein  
A;Reference number: A47233; MUID:93087573; PMID:1454851  
A;Accession: I60906  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-16 <RE2>  
A;Cross-references: GB:M90078; NID:G178786; PIDN:AAA35547.1; PID:G553188  
A;Note: apo(a) Gene 1 (nomenclature of reference 152415)  
A;Accession: A47233  
A;Status: preliminary; translation not shown; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-16 <RE5>  
A;Cross-references: GB:M90079; NID:G178784; PIDN:AAA35546.1; PID:G553187  
R;Ichinose, A.  
Biochemistry 31, 3113-3118, 1992  
A;Title: Multiple members of the plasminogen-apolipoprotein(a) gene family associated with  
A;Reference number: 152415; MUID:92207924; PMID:1554698  
A;Accession: 152415  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-16 <RE3>  
A;Cross-references: GB:M86877; NID:G178780; PIDN:AAB49909.1; PID:G553185  
A;Note: apo(a) Gene 1 (nomenclature of reference 152415)  
A;Accession: I65286  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-16 <RE4>  
A;Cross-references: GB:M86878; NID:G178782; PIDN:AAA51749.1; PID:G553186  
C;Genetics:  
A;Gene: GDB:LPA  
A;Cross-references: GDB:120699; OMIM:152200  
A;Map position: 6q26-6q27  
A;Note: several genes closely linked on chromosome 6 are identical in the first coding exons  
of kringle repeats  
C;Superfamily: apolipoprotein(a); kringle homology; trypsin homology  
C;Keywords: hydrolase; kringle; lipid binding; lipoprotein; serine proteinase  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-4548/Product: apolipoprotein(a) #status experimental <MAT>  
F;28-105/Domain: kringle homology <KR1>  
F;142-219/Domain: kringle homology <KR2>  
F;256-333/Domain: kringle homology <KR3>  
F;370-447/Domain: kringle homology <KR4>  
F;484-561/Domain: kringle homology <KR5>  
F;598-675/Domain: kringle homology <KR6>  
F;712-789/Domain: kringle homology <KR7>  
F;826-903/Domain: kringle homology <KR8>  
F;940-1017/Domain: kringle homology <KR9>  
F;1054-1131/Domain: kringle homology <KR10>  
F;1168-1245/Domain: kringle homology <KR11>  
F;1282-1359/Domain: kringle homology <KR12>  
F;1396-1473/Domain: kringle homology <KR13>  
F;1510-1587/Domain: kringle homology <KR14>  
F;1624-1701/Domain: kringle homology <KR15>  
F;1738-1825/Domain: kringle homology <KR16>  
F;1852-1929/Domain: kringle homology <KR17>  
F;1966-2043/Domain: kringle homology <KR18>  
F;2080-2157/Domain: kringle homology <KR19>  
F;2194-2271/Domain: kringle homology <KR20>  
F;2308-2385/Domain: kringle homology <KR21>  
F;2422-2499/Domain: kringle homology <KR22>  
F;2536-2613/Domain: kringle homology <KR23>  
F;2650-2727/Domain: kringle homology <KR24>  
F;2764-2841/Domain: kringle homology <KR25>  
F;2878-2955/Domain: kringle homology <KR26>  
F;2992-3069/Domain: kringle homology <KR27>  
F;3106-3183/Domain: kringle homology <KR28>  
F;3220-3297/Domain: kringle homology <KR29>  
F;3334-3411/Domain: kringle homology <KR30>  
F;3448-3525/Domain: kringle homology <KR31>  
F;3562-3639/Domain: kringle homology <KR32>  
F;3676-3753/Domain: kringle homology <KR33>

F;3782-3859/Domain: kringle homology <KR34>  
F;3896-3973/Domain: kringle homology <KR35>  
F;4010-4087/Domain: kringle homology <KR36>  
F;4124-4201/Domain: kringle homology <KR37>  
F;4228-4307/Domain: kringle homology <KR38>  
F;4328-4541/Domain: trypsin homology <TRY>  
Query Match 30.6%; Score 156; DB 1; Length 4548;  
Best Local Similarity 37.8%; Pred. No. 2e-08;  
Matches 34; Conservative 12; Mismatches 32; Indels 12; Gaps  
QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPMSATV--LQOTYHAHRSDALQLGLGKHNCRNP 58  
DB 4122 RCYHGNQSYRGTFSTVTGRTCSKSSMTPHRHQRTPEPNPDGLTM-----NYCRNP 41  
QY 59 DNRPRPWCYQVGLKELV--QECMVHDCAD 86  
DB 4177 DADTGPWCFT---MDPSIRWYCNLTRCSD 4203  
RESULT 25  
S45281  
coagulation factor XIIa (EC 3.4.21.38) precursor - bovine (fragment)  
N;Alternate names: Hageman factor (activated)  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 10-Apr-1995 #sequence\_revision 22-Apr-1995 #text\_change 21-Jan-2000  
C;Accession: S45281; A61329  
R;Shibuya, Y.; Semba, U.; Okabe, H.; Kambara, T.; Yamamoto, T.  
Biochim. Biophys. Acta 1206, 63-70, 1994  
A;Title: Primary structure of bovine Hageman factor (blood coagulation factor I): c  
A;Reference number: S45281; MUID:94242782; PMID:8186251  
A;Accession: S45281  
A;Molecule type: mRNA  
A;Residues: 1-593 <SHI>  
A;Cross-references: GB:S70164  
A;Note: the authors translated the codon GAG for residue 23 as Val, GAG for r  
s pro. CTC for residue 203 as Phe, GTG for residue 247 as Leu, CCG for residu  
is, and ATC for residue 505 as Leu  
R;Fujikawa, K.; Walsh, K.A.; Davie, E.W.  
Biochemistry 16, 2270-2278, 1977  
A;Title: Isolation and characterization of bovine factor XII (Hageman factor)  
A;Reference number: A61329; MUID:77182112; PMID:861210  
A;Accession: A61329  
A;Molecule type: protein  
A;Residues: 10-16 'X', 18-19, 525-550 <FUJ>  
C;Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat  
C;Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; monomer;  
F;37-78/Domain: fibronectin type II repeat homology <IF2>  
F;88-120/Domain: EGF homology <EGF>  
F;125-160/Domain: fibronectin type I repeat homology <FB1>  
F;207-287/Domain: kringle homology <KR>  
F;350-587/Domain: trypsin homology <TRY>  
F;541/Active site: Ser #status predicted  
Query Match 30.5%; Score 155.5; DB 2; Length 593;  
Best Local Similarity 38.4%; Pred. No. 3.3e-09;  
Matches 33; Conservative 9; Mismatches 37; Indels 7; Gaps  
QY 2 TCYEGNGHFYRGKASTDTMGRCPLPMSATVLTQTY-HAHRSDALQLGLGKHNCRNP 58  
DB 206 SCYDDRDRLSYRGAGTTLTGAPCQSWAS-----EATYWNVTASQVNLWGLGDHAFCRNP 261  
QY 59 DNRPRPWCYQVGLKPLVQECMVHDC 84  
DB 262 DNDTRPWCFTWKGDRLSNWYCRCLAPC 287  
RESULT 26  
A60140  
plasmin (EC 3.4.21.7) precursor - chicken (fragment)  
N;Alternate names: plasminogen  
C;Species: Gallus gallus (chicken)  
C;Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 16-Jul-1999

A,Cross-references: GB:M62832; NID:g2062155; PIDN:AAA41884.1; PID:g954488  
A>Note: the authors translated the codon TCT for residue 76 as Ala  
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor  
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F:134-112/Domain: Kringle homology <KG>  
F:134-112.55-95.83-107/Disulfide bonds: #status predicted

Query Match 29.3%; Score 149.5; DB 2; Length 169;  
Best Local Similarity 37.8%; Pred. No. 4.6e+09;  
Matches 34; Conservative 12; Mismatches 29; Indels 15; Gaps 6

OY 3 CYEGNGHFYKCASTDTMGPCLEPNSATVLQQTVAHRSDALQL---GLGKHNYCRNPD 59  
DB ||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||  
34 CYQGNGKSYRTGSTTNTGKKCSW-----VSMTPHSHSKTPANFPDSSL-ENMYCRNPD 87

OY 60 N-RRRPWCYVGGLKLPLV--QECMVHDCA 86  
DB ||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||  
88 NDQRGPWCFT---TDPSVRWEYNLKRCE 114

RESULT 29

PUBO

plasmin (EC 3.4.21.7) precursor - bovine

NlAlternate names: plasminogen

C:Species: Bos primigenius taurus (cattle)

C:Date: 30-Sep-1987 #sequence revision 28-Apr-1995 #text\_change 18-Jun-1999

C:Accession: S45046; A25835; I45961; S03736

R:Berglund, L.; Andersen, M.D.; Petersen, T.E.  
submitted to the EMBL Data Library May 1994

A:Description: Cloning and characterization of the bovine plasminogen cDNA.

A:Reference number: S45046

A:Accession: S45046

A:Molecule type: mRNA

A:Residues: 1-812 <BR>

A:Cross-references: EMBL:X79402; NID:g494962; PIDN:CAA55939.1; PID:g494963

A:Experimental source: liver

A>Note: it is uncertain whether Met-1 or Met-8 is the initiator

R:Schaller, J.; Moser, P.W.; Danneberg-Muller, G.A.K.; Rosselet, S.J.; Kampfer  
Eur. J. Biochem. 129, 267-278, 1985

A:title: Complete amino acid sequence of bovine plasminogen. Comparison with h  
man arg, R; Ric

A:Reference number: A25835; MOID:85203906; PMID:3846532

A:Accession: A25835

A:Molecule type: protein

A:Residues: 27-334,'D','336-515','H','517-554','L','556-812 <SCH>

R:Malinowski, D.P.; Sadler, J.E.; Davie, E.W.  
Biochemistry 23, 4243-4250, 1984

A:title: Characterization of a complementary deoxyribonucleic acid coding for  
man arg, I

A:Reference number: I45961; MOID:85023311; PMID:6148961

A:Accession: I45961

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 706-743,'R','745-812 <MAL>

A:Cross-references: GB:K02935; NID:g163551; PIDN:AAA30714.1; PID:g163552

R:Brunisholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manr  
Eur. J. Biochem. 114, 465-470, 1981

A:title: Comparison of the primary structure of the N-terminal CNBr fragments  
human

A:Reference number: S03735; MOID:81212097; PMID:7238497

A:Accession: S03736

A:Molecule type: protein

A:Residues: 27-83 <BRU>

C:Function:

A:Description: dissolves the fibrin of blood clots; acts as a proteolytic fact  
ns the walls of the graafian follicle; also activates the urokinase-type plas  
ogen

A:Pathway: fibrinolysis

C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor  
C:Keywords: duplication; fibrinolysis; glycoprotein; hydrolase; kidney; kringl  
plasi

F:1-26/Domain: signal sequence #status predicted <SIG>

F:8-103/Domain: plasminogen-related protein precursor homology <PluPH>

E:27-812/Product: plasminogen #status experimental <PRO>

F:27-103/Domain: activation peptide #status experimental <APT>

F:104-583,584-812/Product: plasmin #status experimental <MA>

F:104-583/Domain: plasmin chain A #status experimental <ACH>

F:110-188/Domain: plasminogen homology <PR1>

A;Accession: S28200  
A:Molecule type: protein  
A:Residues: 118-460 <Sc2>  
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homologous Cysteine protease; zymogens; fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase; Zymogen family  
F:1-37/Domain: activation peptide (fragment) #status experimental <AP>  
F:38-117/118-230/231-460/Product: plasmin (fragments) #status experimental <MAT>  
F:41-118/Domain: kringle homology <KR4>  
F:418-460/Product: miniplasminogen #status experimental <MIN>  
F:432-211/Domain: kringle homology <KR5>  
F:226-460/Domain: plasmin chain B #status experimental <BCH>  
F:231-453/Domain: trypsin homology <TRY>  
F:272,315,410/Active site: His, Asp, Ser #status predicted

Query Match            28.9%; Score 147.5; DB 2; Length 460;  
Best Local Similarity 43.3%; Pred. No. 2e-08;  
Matches 29; Conservative 8; Mismatches 23; Indels 7; Gaps 2;

Qy     3 CYEGNGHYRGKASTDTMGRLPWSNATV-LQQTYHAHRSDALGLGCKNYCRNPNN 60  
     | ||| :||| :||| :||| :||| :  
Db     41 CYHGNGGVYRGTSTTVTKGRKSQSMSIPRRHQKTPEYPNAGLTM----NYCRNPDA 95  
  
Qy     61 RRPWPWCY 67  
     : ||| :  
Db     96 DKSPWCV 102

RESULT 33  
B45082  
N:Contains: receptor ror2 precursor - human neurotrophic protein-tyrosine kinase (EC 2.7.1.112)  
C:Species: Homo sapiens (man)  
C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 19-May-2000  
C:Accession: B45082  
J:Maslakowski, P.; Carroll, R.D.  
U:Biochemistry, 267, 26181-26190, 1992  
A>Title: A novel family of cell surface receptors with tyrosine kinase-like domain.  
A:Reference number: A45082; PMID:93100347; PMID:1334494  
A:Accession: B45082  
A:Molecule type: mRNA  
A:Residues: 1-943 <MAS>  
A:CROSS-references: GB:M97639; NID:g337466; PIDN:AAA60276.1; PID:g337467  
A>Note: sequence extracted from NCBI backbone (NCBIB:120918)  
C:Genetics:  
A:Gene: GDB:NTRKR2  
A:CROSS-references: GDB:136454  
A:Map position: 6p21-6p21  
C:Superfamily: neurotrophic receptor ror; immunoglobulin homology; kringle homology; prokineticin receptor; tyrosine phosphorylation; transmembrane protein; tyrosine phosphatase  
F:1-27/Domain: signal sequence #status predicted <SIG>  
F:28-937/Product: neurotrophic receptor ror2 #status predicted <MAT>  
F:16-137/Domain: immunoglobulin homology <IMM>  
F:316-394/Domain: kringle homology <KR>  
F:412-428/Domain: transmembrane #status predicted <TMN>  
F:471-753/Domain: protein kinase homology <KIN>  
F:479-487/Region: protein kinase ATP-binding motif  
F:770,188,318/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match            28.9%; Score 147.5; DB 2; Length 943;  
Best Local Similarity 39.3%; Pred. No. 3.8e-08;  
Matches 33; Conservative 10; Mismatches 36; Indels 5; Gaps 4;

Qy     3 CYEGNGHYRGKASTDTMGRLPWSNATVLQQTYHAHRSDALGLGCKNYCRNPNNR 62  
     | ||| :||| :||| :||| :||| :  
Db     316 CYNSSGMRYRTATTAKSGHCQPW--ALQHPSHLHSLDTPFLG-GGHAYCNFGQM 372  
  
Qy     63 R-PWCYVGQLKLPLVECMWFDC 85  
     : ||| :  
Db     373 EGPWCFTQ-NKNVRMELCDVPSCS 395

RESULT 34

```

F.192-269/Domain: kringle homology <KR2>
F.282-359/Domain: kringle homology <KR3>
F.384-461/Domain: kringle homology <KR4>
F.485-564/Domain: kringle homology <KR5>
F.584-812/Domain: plasmin chain B #status experimental <BCH>
F.584-805/Domain: trypsin homology <TRY>
F.56-80,60-68,110-188,131-171,159-183,192-269,195-323,213-252,241-264,282-359,303-342,33-
bonds: #status predicted
F.315/Binding site: Carbohydrate (Asn) (covalent) #status experimental
F.365/Binding site: Carbohydrate (Ser) (covalent) #status experimental
F.624,667,762/Active site: His, Asp, Ser #status predicted

Query Match          29.2%;   Score 149;   DB 1;   Length 812;
Best Local Similarity 39.3%;   Pred. No. 2.3e-08;
Matches 35;   Conservative 8;   Mismatches 32;   Indels 14;   Gaps 5;

QY      3  CYEGNGHFYRGKASDTDMRGCLPWNSAIVLQTYHAH---RSDALQLGKHNCRND 59
Ddb     384  CYHNGQSVRGHSSTITIGKQSWSS-----MTPRHUKTPTNPVAGL-TWNYCRND 437

QY      60  NRRRPWCYVQVGLKPLV--QECWHDCA  86
Ddb     438  ADKSPWCYT---TDPRVRWEFCNLKCKSE  463

```

RESULT 30  
I46260  
precursor - western European hedgehog  
C.Species: Erinaceus europaeus (western European hedgehog)  
C.Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 16-Jul-1999  
C.Accession: I46260  
J. Biol. Chem. 270, 24004-24009, 1995  
A.Title: The recurring evolution of Lp(a): Insights from cloning of hedgehog apolipoprotein A1  
A.Reference number: I46259; MUID:96025778; PMID:7592597  
A.Accession: I46260  
A.Status: preliminary; translated from GB/EMBL/DBJ  
A.Molecule type: mRNA  
A.Residues: 1-810 <LAW>  
A.Cross-references: EMBL:U33171; NID:G1046360; PID:G1046361  
C.Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology  
C.Keywords: hydrolase; serine proteinase  
F.1-96/Domain: plasminogen-related protein precursor homology <PLPH>  
F.103-181/Domain: kringle homology <KR1>  
F.185-262/Domain: kringle homology <KR2>  
F.273-352/Domain: kringle homology <KR3>  
F.379-456/Domain: kringle homology <KR4>  
F.482-561/Domain: kringle homology <KR5>  
F.582-803/Domain: trypsin homology <TRY>

```

Query Match          29.1%; Score 148.5; DB 2; Length 810;
Best Local Similarity 41.5%; Pred. No. 2.6e-08;
Matches 27; Conservative 7; Mismatches 28; Indels 3; Gaps 1;

3 CYEENGHFVRGKASTDTMRPCLPWNSATVLQQTYHAHRSDALQLGLGKNYCRNPDNR 52
|||||:|||||:|||||:|||||:|||||:
379 CYQNGQIYRGTSSTITGKKQPWTSMRPHRSHKTPENYPDADLTM---NYCRNPDGDK 435
|||||:|||||:|||||:|||||:|||||:

QY 63 RPWCY 67
|||||
436 GPWCY 440

RESULT 31
PLUG
plasmin (EC 3.4.21.7) precursor - pig (fragment)
N1Alternate names: Plasminogen
N1Contains: miniplasminogen
C1Species: Sus scrofa domestica (domestic pig)
C1Date: 07-Sep-1990 #sequence revision 01-Nov-1996 #text_change 18-Jul-1997
C1Accession: S03733; S03737; A25834
C1R1Schaller, J.; Marti, T.; Roeseleat, S.J.; Kaempfer, U.; Rickli, E.E.
C1Brincolysis 1, 91-102, 1987

```

A:Title: Amino acid sequence of the heavy chain of porcine plasmin. Comparison f the  
A:Reference number: S03733  
A:Accession: S03733  
A:Molecule type: protein  
A:Residues: 1-560 <SCH>  
R:Brunisholz, R.A.; Lersch, P.G.; Schaller, J.; Rickli, E.E.; Lergler, W.; Mann  
Eur. J. Biochem. 114, 465-470, 1981  
A:Title: Comparison of the primary structure of the N-terminal CNBr fragments  
A:Reference number: S03735; MUID:8121097; PMID:7238497  
A:Accession: S03737  
A:Molecule type: protein  
A:Residues: 1-57 <BRU>  
R:R.Marti, T.; Schaller, J.; Rickli, E.E.  
Eur. J. Biochem. 149, 279-285, 1985  
A:Title: Determination of the complete amino-acid sequence of porcine miniplas  
A:Reference number: A25834; MUID:85203907; PMID:3846533  
A:Accession: A25834  
A:Molecule type: protein  
A:Residues: 450-790 <VAR>  
C:Function:  
A:Description: dissolves the fibrin of blood clots; acts as a proteolytic fact in a  
ns the walls of the graafian follicle; also activates the urokinase-type plas  
A:Pathway: fibrinolysis  
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor  
C:Keywords: fibrinolysis; kringle homology; hydrolase; kidney; kringle; plasma; s  
F:1-790/Product: plasminogen #status predicted <PRO>  
F:1-77/Domain: plasminogen-related protein precursor homology (fragment) <PLPE  
F:78-560/Product: plasmin chain A #status predicted <APT>  
F:84-162/Domain: kringle homology <KR1>  
F:166-243/Domain: kringle homology <KR2>  
F:256-333/Domain: kringle homology <KR3>  
F:358-435/Domain: kringle homology <KR4>  
F:450-790/Product: miniplasminogen #status experimental <MIN>  
F:461-540/Domain: kringle homology <KR5>  
F:561-790/Product: plasmin chain B #status experimental <BCH>  
F:561-783/Domain: trypsin homology <TRY>  
F:30-54, 34-42, 84-162, 105-145, 133-157, 166-243, 169-297, 187-226, 215-238, 256-333, 2  
bonds: #status predicted  
602.645, 740/Active site: His, Asp, Ser #status predicted

```

Query Match          29.0%; Score 148; DB 1; Length 790;
Best Local Similarity 38.2%; Pred. No. 2.9e-08;
Matches 34; Conservative 8; Mismatches 33; Indels 14; Gaps 5

QY      3 CYEGNGHYRGKASTDTMGRCLPWNSATVLOQTYHAHR-----SDALQLGLGKHNYCRNPD 59
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      358 CYRGNGESYRGTSSTTIGRKQCSW-----VSMTPHRHEKTPGNFPNAGL-TMNYCRNPD 411
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      60 NRRFPWCYVQGLXPLV--QECMVHDCAD 86
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      412 ADKSPWCYT--TDPVRVWEYCNLKGCS 437
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 32
B61545
plasmin (EC 3.4.21.7) precursor - sheep (fragments)
N:Alternate names: plasminogen
N:Contains: miniplasminogen
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 28-Oct-1994 #sequence_revision 01-Nov-1996 #text_change 17-Mar-1999
C:Accession: B61545; S28200
R:Schaller, J.; Rickli, E.E.
Enzyme 40.63-69, 1988

A:Title: Structural aspects of the plasminogen of various species.
A:Reference number: A61545; MUID:89005015; PMID:3188975
A:Accession: B61545
A:Molecule type: protein
A:Residues: 1-37;38-117 <SCH>
R:Schaller, J.; Straub, C.; Kaempfer, U.; Rickli, E.E.
Protein Seq. Data Anal. 5, 21-25, 1992
A:Title: Complete amino acid sequence of ovine miniplasminogen.
A:Reference number: S28200; MUID:93149395; PMID:1492092

```



A:Molecule type: mRNA  
A:Residues: 1-471,'D',473-810 <FOR>  
A:Cross-references: GB:X05199; NID:G35530; PIDN:CAA28831.1; PID:G35531  
A:Experimental source: liver  
R:Malinowski, D.P.; Sadler, J.E.; Davie, E.W.  
Biochemistry 23, 4243-4250, 1984  
A:Title: Characterization of a complementary deoxyribonucleic acid coding for human and  
A:Reference number: I45966; MUID:85023311; PMID:6148961  
A:Accession: I62738  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 292-471,'D',473-810 <MAL2>  
A:Cross-references: GB:K02922; NID:G190112; PIDN:AAA60124.1; PID:G387031  
A:Experimental source: liver  
R:Malinowski, D.P.; Sadler, J.E.; Davie, E.W.  
Biochemistry 23, 4243-4250, 1984  
A:Title: Characterization of a complementary deoxyribonucleic acid coding for human and  
A:Reference number: I45966; MUID:85023311; PMID:6148961  
A:Accession: I62738  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 367-419 <MAL3>  
A:Cross-references: GB:K02921; NID:G190110; PIDN:AAA60123.1; PID:G190111  
R:Brumsholtz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M.;  
Eur. J. Biochem. 114, 465-470, 1981  
A:Title: Comparison of the primary structure of the N-terminal CNBr fragments of human,  
A:Reference number: S03735; MUID:81212097; PMID:7238497  
A:Accession: S03735  
A:Molecule type: protein  
A:Residues: 20-71,'E',73-76 <BRU>  
R:Sottrop-Jensen, L.; Petersen, T.E.; Magnusson, S.  
Submitted to the Atlas, July 1977  
A:Reference number: A00929  
A:Accession: A00929  
A:Molecule type: protein  
A:Residues: 20-71,'E',73-85,87-106,'D',108-360,'E',362-810 <SOT>  
R:Wiman, B.  
Eur. J. Biochem. 76, 129-137, 1977  
A:Title: Primary structure of the B-chain of human plasmin.  
A:Reference number: A04627; MUID:7725245; PMID:42009  
A:Accession: A04627  
A:Molecule type: protein  
A:Residues: 581-810 <W1>  
R:Wiman, B.; Wallen, P.  
Eur. J. Biochem. 50, 489-494, 1975  
A:Title: Structural relationship between "glutamic acid" and "lysine" forms of human plasmin  
A:Reference number: A04625; MUID:75093329; PMID:122932  
A:Accession: A04625  
A:Molecule type: protein  
A:Residues: 20-50,'Q',51-71,'E',73-85,87-100 <W12>  
R:Wiman, B.; Wallen, P.  
Eur. J. Biochem. 58, 539-547, 1975  
A:Title: Amino-acid sequence of the cyanogen-bromide fragment from human plasminogen the  
A:Reference number: A04626; MUID:76043692; PMID:126863  
A:Accession: A04626  
A:Molecule type: protein  
A:Residues: 483-507,'E',509-604 <W13>  
R:Robbins, K.C.; Bernabe, P.; Arzadon, L.; Summaria, L.  
J. Biol. Chem. 248, 1631-1633, 1973  
A:Title: The primary structure of human plasminogen. II. The histidine loop of human plasminogen  
A:Reference number: A92125; MUID:73149248; PMID:4694729  
A:Contents: annotation; active site  
R:Groskopf, W.R.; Summaria, L.; Robbins, K.C.  
J. Biol. Chem. 244, 3530-3597, 1969  
A:Title: Studies on the active center of human plasmin. Partial amino acid sequence of a  
A:Reference number: A92048; MUID:69234739; PMID:4240117  
A:Contents: annotation; active site  
R:Trexler, M.; Vali, Z.; Pathy, L.  
J. Biol. Chem. 257, 7401-7406, 1982  
A:Title: Structure of the omega-aminocarboxylic acid-binding sites of human plasminogen.  
A:Reference number: A92382; MUID:82213905; PMID:6919539  
A:Contents: annotation; omega-aminocarboxylic acid binding sites  
R:Vali, Z.; Pathy, L.  
J. Biol. Chem. 259, 13690-13694, 1984  
A:Title: The fibrin-binding site of human plasminogen. Arginines 32 and 34 are essential  
A:Reference number: A92458; MUID:85054794; PMID:6094526  
A:Contents: annotation; fibrin binding site; omega-aminocarboxylic acid binding site

R:Gao, Y.; Ji, R.W.; Davidson, D.; Schaller, J.; Marti, D.; Soehndel, S.; McCr e, S.  
J. Biol. Chem. 271, 29461-29467, 1996  
A:Title: Kringle domains of human angiotensin. Characterization of the anti-pl  
A:Reference number: A58811; MUID:97087211; PMID:8910613  
A:Contents: annotation  
R:Ljajnen, H.R.; Ugwu, F.; Bini, A.; Collen, D.  
Biochemistry 37, 4699-4702, 1998  
A:Title: Generation of an angiotensin-like fragment from plasminogen by strom  
A:Reference number: A58812; MUID:9548733; PMID:9548733  
A:Contents: annotation  
R:Tulinsky, A.; Mulichak, A.M.  
submitted to the Brookhaven Protein Data Bank, July 1991  
A:Reference number: A51341; PDB:1PK4  
A:Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 376-45  
R:Tulinsky, A.; Wu, T.P.  
submitted to the Brookhaven Protein Data Bank, July 1991  
A:Reference number: A51488; PDB:2PK4  
A:Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 375-4  
R:Wu, T.P.; Tulinsky, A.  
submitted to the Brookhaven Protein Data Bank, August 1993  
A:Reference number: A51911; PDB:1PKR  
A:Contents: annotation; X-ray crystallography, 2.48 angstroms, residues 102-1  
R:Padmanabhan, K.; Tulinsky, A.  
submitted to the Brookhaven Protein Data Bank, April 1994  
A:Reference number: A52408; PDB:1PMK  
A:Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 377-4  
R:Tulinsky, A.; Mathews, I.I.  
submitted to the Brookhaven Protein Data Bank, December 1995  
A:Reference number: A65244; PDB:1CEA  
A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-18  
R:Tulinsky, A.; Mathews, I.I.  
submitted to the Brookhaven Protein Data Bank, December 1995  
A:Reference number: A65245; PDB:1CEB  
A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-18  
R:Mulichak, A.M.; Tulinsky, A.; Ravichandran, K.G.  
Biochemistry 30, 10576-10588, 1991  
A:Title: Crystal and molecular structure of human plasminogen kringle 4 refine  
A:Reference number: A58819; MUID:92031502; PMID:1657148  
A:Contents: annotation  
R:Wu, T.P.; Padmanabhan, K.; Tulinsky, A.; Mulichak, A.M.  
Biochemistry 30, 10589-10594, 1991  
A:Title: The refined structure of the epsilon-aminocaproic acid complex of hu  
A:Reference number: A58818; MUID:92031503; PMID:1657149  
A:Contents: annotation  
R:de Vos, A.M.; Ultsch, M.H.; Kelley, R.F.; Padmanabhan, K.; Tulinsky, A.; Wes  
Biochemistry 31, 270-279, 1992  
A:Title: Crystal structure of the kringle 2 domain of tissue plasminogen actin  
A:Reference number: A39483; MUID:9211803; PMID:1310033  
A:Contents: annotation; X-ray crystallography, 2.4 angstroms  
R:Stec, B.; Teeter, M.M.; Whitlow, M.; Yamano, A.  
submitted to the Brookhaven Protein Data Bank, June 1995  
A:Reference number: A65980; PDB:1KRN  
A:Contents: annotation; X-ray crystallography, 1.67 angstroms, residues 376-4  
R:Rejante, M.; Llinas, M.  
submitted to the Brookhaven Protein Data Bank, August 1996  
A:Reference number: A65803; PDB:1HPJ  
A:Contents: annotation; conformation by (1)H-NMR, residues 103-181  
R:Rejante, M.; Llinas, M.  
submitted to the Brookhaven Protein Data Bank, August 1996  
A:Reference number: A65804; PDB:1HPK  
A:Contents: annotation; conformation by (1)H-NMR, residues 103-181  
R:Rejante, M.R.; Llinas, M.  
Eur. J. Biochem. 221, 927-937, 1994  
A:Title: (1)H-NMR assignments and secondary structure of human plasminogen kr  
A:Reference number: A58817; MUID:94237157; PMID:8181475  
R:Rejante, M.R.; Llinas, M.  
conformation by (1)H-NMR, residues 96-184  
R:Rejante, M.R.; Llinas, M.  
Eur. J. Biochem. 221, 939-949, 1994  
A:Title: Solution structure of the epsilon-aminohexanoic acid complex of huma  
A:Reference number: A58817; MUID:94237158; PMID:8181476  
A:Contents: annotation; conformation by (1)H-NMR  
C:Comment: Plasminogen is synthesized by the kidney and is present in plasma  
C:Comment: Plasminogen is converted to plasmin by plasminogen activators (see





[illegible]

A:Molecule type: DNA  
A:Residues: 1-30 <RES>  
A:Cross-references: ENBL:X81630; NID:G673451; PIDN:CAA57286.1; PID:G673452  
C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor  
C:Function:  
A:Description: stimulates mitosis of hepatocytes and other cells  
A:Note: does not have proteinase activity  
C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology  
C:Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringle; pyro  
F:1-32/Domain: signal sequence #status predicted <SIG>  
F:56-495/Domain: signal sequence #status predicted <SIG>  
F:56-495/Domain: hepatocyte growth factor #status predicted <MAT>  
F:129-207/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>  
F:129-207/Domain: kringle homology <KR1>  
F:129-207/Domain: kringle homology <KR2>  
F:306-384/Domain: kringle homology <KR3>  
F:392-470/Domain: kringle homology <KR4>  
F:496-728/Domain: hepatocyte growth factor beta chain #status predicted <BCH>  
F:496-728/Domain: trypsin homology <TRY>  
F:33/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted  
F:295,403,569,656/Binding site: carboxylate (Asn) (covalent) #status predicted  
F:488-607/Disulfide bonds: #status predicted

Query Match 27.0%; Score 137.5; DB 1; Length 728;  
Best Local Similarity 32.0%; Pred. No. 3.7e-07; Mismatches 12; Indels 19; Gaps 4;  
Matches 31; Conservative 12; Mismatches 35; Indels 19; Gaps 4;  
QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOYHAHRSDALQGLG----KHNYCRNP 58  
DB 306 CIQGGEGYRGNTIWIWGIPQWRDS-----QYPKHHDITPENFKCKDLRENYCRNP 358  
QY 59 DNRPRPCV-----VOVGLKPLVQECMV---HDCADG 87  
DB 359 DGAESPWCFTTDPNIRGVCSQIPKCDVSSGQDCYRG 395

RESULT 41  
A35644  
N:Alternate names: hepatocyte growth factor precursor - rat  
N:Alternate names: hepatocyte growth factor precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 28-Sep-1990 #sequence\_revision 18-Nov-1992 #text\_change 21-Jul-2000  
C:Accession: A35644; S13211  
R:Tashiro, K.; Hagiya, M.; Nishizawa, T.; Seki, T.; Shimomishi, M.; Shimizu, S.; Nakamura  
Proc. Natl. Acad. Sci. U.S.A. 87, 3200-3204, 1990  
A:Title: Deduced primary structure of rat hepatocyte growth factor and expression of the  
A:Reference number: A35644; MUID:90222197; PMID:2139229  
A:Accession: A35644  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-728 <TAS>  
A:Cross-references: GB:D90102; GB:M32987; NID:G220766; PIDN:BAA14133.1; PID:G220767  
A:Note: the authors translated the codon GAG for residue 70 as Gln, GAC for residue 417  
R:Okajima, A.; Miyazawa, K.; Kitamura, N.  
Eur. J. Biochem. 193, 375-381, 1990  
A:Title: Primary structure of rat hepatocyte growth factor and induction of its mRNA dur  
A:Reference number: S13211; MUID:91031482; PMID:2146117  
A:Accession: S13211  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-728 <OKA>  
A:Cross-references: ENBL:X54400; NID:G56353; PIDN:CAA39266.1; PID:G4539554  
C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor  
C:Function:  
A:Description: stimulates mitosis of hepatocytes and other cells  
A:Note: does not have proteinase activity  
C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology  
C:Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringle; pyro  
F:1-32/Domain: signal sequence #status predicted <SIG>  
F:56-495/Domain: signal sequence #status predicted <SIG>  
F:56-495/Domain: hepatocyte growth factor #status predicted <MAT>  
F:129-207/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>  
F:129-207/Domain: kringle homology <KR1>  
F:129-207/Domain: kringle homology <KR2>  
F:306-384/Domain: kringle homology <KR3>

F:392-470/Domain: kringle homology <KR4>  
F:496-728/Domain: hepatocyte growth factor beta chain #status predicted <BCH>  
F:496-728/Domain: trypsin homology <TRY>  
F:33/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status  
F:295,403,569,656/Binding site: carboxylate (Asn) (covalent) #status predicted  
F:488-607/Disulfide bonds: #status predicted

Query Match 26.6%; Score 135.5; DB 1; Length 728;  
Best Local Similarity 32.0%; Pred. No. 6.2e-07; Mismatches 36; Indels 19; Gaps 4;  
Matches 31; Conservative 11; Mismatches 36; Indels 19; Gaps 4;  
QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOYHAHRSDALQGLG----KHNYCRNP 58  
DB 306 CIQGGEGYRGNTIWIWGIPQWRDS-----QYPKHHDITPENFKCKDLRENYCRNP 358  
QY 59 DNRPRPCV-----VOVGLKPLVQECMV---HDCADG 87  
DB 359 DGAESPWCFTTDPNIRGVCSQIPKCDVSSGQDCYRG 395

RESULT 42  
A61545  
Plasmin (EC 3.4.21.7) precursor - horse (fragments)  
N:Alternate names: plasminogen  
N:Contains: miniplasminogen  
C:Species: Equus caballus (domestic horse)  
C:Date: 28-Oct-1994 #sequence\_revision 01-Nov-1996 #text\_change 18-Jul-1997  
C:Accession: A61545; S17527  
R:Schaller, J.; Rickli, E.E.  
Enzyme 40, 63-69, 1988  
A:Title: Structural aspects of the plasminogen of various species.  
A:Reference number: A61545; MUID:89005015; PMID:3168975  
A:Accession: A61545  
A:Molecule type: protein  
A:Residues: 1-33;34-117 <SCH>  
R:Schaller, J.; Straub, C.; Kaempfer, U.; Rickli, E.E.  
Protein Seq. Data Anal. 4, 69-74, 1991  
A:Title: Complete amino acid sequence of equine miniplasminogen.  
A:Reference number: S17527; MUID:92052077; PMID:1946332  
A:Accession: S17527  
A:Molecule type: protein  
A:Residues: 118-455 <SC2>  
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor  
C:Keywords: fibrinolysis; glycoprotein; hyaluronase; kringle; plasmin; serine pro  
F:1-33;34-117;118-455/Product: plasminogen (fragments) #status experimental <P  
F:1-33/Domain: activation peptide (fragment) #status experimental <APT>  
F:34-117,118-225,226-455/Product: plasmin (fragments) #status experimental <MA  
F:37-114/Domain: kringle homology <KR4>  
F:118-455/Product: miniplasminogen #status experimental <MIN>  
F:126-205/Domain: kringle homology <KR5>  
F:226-455/Domain: plasmin chain B #status experimental <BCH>  
F:226-448/Domain: trypsin homology <TRY>  
F:267,310,405/Active site: His, Asp, Ser #status predicted

Query Match 26.3%; Score 134; DB 2; Length 455;  
Best Local Similarity 34.1%; Pred. No. 5.8e-07;  
Matches 30; Conservative 13; Mismatches 33; Indels 12; Gaps 4  
QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATV--LQOYHAHRSDALQGLGHNYCRNP 60  
DB 37 CYQKGSYRGTSITVTGKKCSWSMTPHWHQKTPKYPNADLTM-----NYCRNPDG 91  
QY 61 RRRPWCYVQVGLKPLV--QECMVHDCAD 86  
DB 92 DKGPWCYT---TDPVSRWEFCNLKCS 116

RESULT 43  
JH0579  
hepatocyte growth factor precursor [validated] - human  
N:Alternate names: hepatopietin A; scatter factor  
C:Species: Homo sapiens (man)  
C:Date: 17-Aug-1992 #sequence\_revision 17-Aug-1992 #text\_change 08-Dec-2000

A;Accession: JH0579; J03033; A41140; B36677; A36677; A33512; A39006; PH0114; A37796; S06  
R;Seki, T.; Hagiya, M.; Shimomishi, M.; Nakamura, T.; Shimizu, S.  
Gene 102, 213-219, 1991  
A;Title: Organization of the human hepatocyte growth factor-encoding gene.  
A;Reference number: JH0579; MUID:91340155; PMID:1831432  
A;Accession: JH0579  
A;Molecule type: DNA  
A;Residues: 1-728 <SEK>  
A;Cross-references: DBJ:J03033  
A;Note: The authors translated the codon GAA for residue 662 as Gly  
R;Seki, T.; Hagiya, M.; Shimomishi, M.; Nakamura, T.; Shimizu, S.  
submitted to JIPID, March 1991  
A;Description: Organization of the human hepatocyte growth factor-encoding gene.  
A;Reference number: J03033  
A;Accession: J03033  
A;Molecule type: DNA  
A;Residues: 1-481, RT, 484-728 <SE2>  
R;Weidner, K.M.; Arakaki, N.; Hartmann, G.; Vandekerckhove, J.; Weingart, S.; Rieder, H.  
Proc. Natl. Acad. Sci. U.S.A. 88, 7001-7005, 1991  
A;Title: Evidence for the identity of human scatter factor and human hepatocyte growth factor.  
A;Reference number: A41140; MUID:91334393; PMID:1831266  
A;Accession: A41140  
A;Molecule type: mRNA  
A;Residues: 1-728 <WE1>  
A;Cross-references: GB:M73239; NID:G337935; PIDN:AAA64239.1; PID:G337936  
R;Seki, T.; Ihara, I.; Sugimura, A.; Shimomishi, M.; Nishizawa, T.; Asami, O.; Hagiya, M.  
Biochem. Biophys. Res. Commun. 172, 321-327, 1990  
A;Title: Isolation and expression of cDNA for different forms of hepatocyte growth factor.  
A;Reference number: A36677; MUID:91025062; PMID:2145836  
A;Accession: A36677  
A;Molecule type: mRNA  
A;Residues: 1-728 <SE3>  
A;Cross-references: GB:M60718; NID:G184031; PIDN:AAA52648.1; PID:G184032  
A;Accession: A36677  
A;Molecule type: mRNA  
A;Residues: 1-161, 167-728 <SE4>  
A;Cross-references: EMBL:X16323  
A;Experimental source: leukocyte  
R;Miyazawa, K.; Tsubouchi, H.; Naka, D.; Takahashi, K.; Okigaki, M.; Arakaki, N.; Nakaya  
Biochem. Biophys. Res. Commun. 163, 967-973, 1989  
A;Title: Molecular cloning and sequence analysis of cDNA for human hepatocyte growth fac  
A;Reference number: A33512; MUID:89392017; PMID:2528952  
A;Accession: A33512  
A;Status: not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-728 <MIY>  
A;Cross-references: GB:M29145; NID:G184041; PIDN:AAA52650.1; PID:G306846  
R;Rubin, J.S.; Chan, A.M.L.; Bottaro, D.P.; Burgess, W.H.; Taylor, W.G.; Cech, A.C.; Hir  
Proc. Natl. Acad. Sci. U.S.A. 88, 415-419, 1991  
A;Title: A broad-spectrum human lung fibroblast-derived mitogen is a variant of hepatocyte  
A;Reference number: A39006; MUID:9110540; PMID:1824873  
A;Accession: A39006  
A;Molecule type: mRNA  
A;Residues: 1-161, 167-728 <RUB>  
A;Cross-references: GB:M35379  
A;Experimental source: embryonic lung  
R;Yoshiyama, Y.; Arakaki, N.; Naka, D.; Takahashi, K.; Hirono, S.; Kondo, J.; Nakayama,  
Biochem. Biophys. Res. Commun. 175, 660-667, 1991  
A;Title: Identification of the N-terminal residue of the heavy chain of both native and  
A;Reference number: PH0114; MUID:91207365; PMID:1826837  
A;Accession: PH0114  
A;Molecule type: protein  
A;Residues: 32-43; 53-58 <YOS>  
A;Experimental source: plasma  
R;Weidner, K.M.; Behrens, J.; Vandekerckhove, J.; Birchmeier, W.  
J. Cell Biol. 111, 2097-2108, 1990  
A;Title: Scatter factor: molecular characteristics and effect on the invasiveness of epi  
A;Reference number: A37796; MUID:91035621; PMID:2146276  
A;Accession: A37796  
A;Molecule type: protein  
A;Residues: 86-91; 329-344; 356-363, 'XX', 366-370; 425-434; 442-447, 'X', 449-450; 543-546, 'X', 5  
R;Nakamura, T.; Nishizawa, T.; Hagiya, M.; Seki, T.; Shimomishi, M.; Sugimura, A.; Tashi  
Nature 342, 440-443, 1989

A;Title: Molecular cloning and expression of human hepatocyte growth factor.  
A;Reference number: S06794; MUID:90066676; PMID:2531289  
A;Accession: S06794  
A;Molecule type: mRNA  
A;Residues: 1-31, 'HK', 34-77, 'N', 79-292, 'V', 294-299, 'M', 301-316, 'A', 318-335, 'K', 337-386  
A;Cross-references: EMBL:X16323; NID:G32081; PIDN:CAA34387.1; PID:G32082  
A;Experimental source: liver  
A;Note: the authors translated the codon CAG for residue 727 as Glu  
R;Hartmann, G.; Naldini, L.; Weidner, K.M.; Sachs, M.; Vigna, E.; Comoglio, P.  
Proc. Natl. Acad. Sci. U.S.A. 89, 11574-11578, 1992  
A;Title: A functional domain in the heavy chain of scatter factor/hepatocyte growth factor.  
A;Reference number: I59214; MUID:93067571; PMID:1280830  
A;Accession: I59214  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-288, 'ET', <HAR>  
R;Miyazawa, K.; Kitamura, G.; Naka, D.; Kitamura, N.  
Eur. J. Biochem. 197, 15-22, 1991  
A;Title: An alternatively processed mRNA generated from human hepatocyte growth factor.  
A;Reference number: S15443; MUID:91200041; PMID:1826653  
A;Accession: S15443  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-288, 'ET', <MIY2>  
A;Cross-references: EMBL:X57574; NID:G32083; PIDN:CAA40802.1; PID:G32084  
R;Shima, N.; Nagao, M.; Osaki, F.; Tsuda, E.; Murakami, A.; Higashio, K.  
Biochem. Biophys. Res. Commun. 180, 1151-1158, 1991  
A;Title: Tumor cytotoxic factor/hepatocyte growth factor from human fibroblast.  
A;Reference number: I52253; MUID:92062058; PMID:1835383  
A;Accession: I52253  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 161-166 <SHI>  
A;Cross-references: GB:S62561; NID:G237996; PIDN:AAB20169.1; PID:G237997  
A;Accession: GDB:RGF  
A;Gene: GDB:RGF  
A;Cross-references: GDB:127524; OMIM:142409  
A;Map position: 7q21.1-7q21.1  
A;Introns: 30/1; 85/2; 123/1; 161/2; 209/1; 249/2; 289/1; 347/2; 390/1; 424/2;  
A;Complex: disulfide-bonded heterodimer of chains derived from the same precursor.  
C;Function:  
A;Description: stimulates mitosis of hepatocytes and other cells  
A;Note: does not have proteinase activity  
C;Superfamily: hepatocyte growth factor; kringle homology; trypsin homology  
C;Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringle; p  
F;1-31/Domain: signal sequence #status predicted <SIG>  
F;32-494/Domain: product hepatocyte growth factor #status experimental <MAT>  
F;32-494/Domain: alpha chain #status experimental <ACH>  
F;128-206/Domain: kringle homology <KR1>  
F;211-288/Domain: kringle homology <KR2>  
F;305-383/Domain: kringle homology <KR3>  
F;391-469/Domain: kringle homology <KR4>  
F;495-728/Domain: beta chain #status experimental <BCH>  
F;495-716/Domain: trypsin homology <TRY>  
F;32/Modified site: pyroldone carboxylic acid (Gln) (in mature form) #status  
F;294,402,566,653/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;494-604/Disulfide bonds: #status predicted

Query Match 26.2%; Score 133.5; DB 1; Length 728;  
Best Local Similarity 34.7%; Pred. No. 1e-06;

Matches 33; Conservative 10; Mismatches 37; Indels 15; Gaps 5

QY 3 CYEKGHFYRGKASTDTMGRCPLPMSATVLOQTYAH--RSDALQLGKKNYCRNPDN 60

Db 305 CIOQGEYRGVTVNTWNGIPQRMDS----QYPHEHDMTFNFKDLRENYCRNPDG 355

QY 61 RREPWCY----VQVQKPLVQEC-MVH--DCADG 87

Db 360 SESFWCTTDPNIRVGYCSQIPNCDMSHQDCYRG 394

RESULT 46  
A40332  
macrophage-stimulating protein 1 precursor - mouse  
NAlternate names: hepatocyte growth factor-like protein  
CSpecies: Mus musculus (house mouse)  
CDate: 17-Jul-1992 #sequence  
CAccession: A40332; B4032  
R'Degen, S.J.F.; Stuart, L.A.; Han, S.; Jamison, C.S.  
#text\_change 18-Jun-1999

ns the walls of the graafian follicle; also activates the urokinase-type plasminogen act  
A:Pathway: fibrinolysis  
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology  
C:Keywords: angiotensin inhibitor; blood; duplication; fibrinolysis; glycoprotein; hyd  
F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-812/Product: plasminogen #status predicted <PRO>  
F:20-96/Domain: activation peptide #status predicted <APT>  
F:49-466/Product: angiotensin #status predicted <AST>  
F:97-581-582-812/Product: plasmin #status predicted <MAT>  
F:97-581-582-812/Product: plasmin #status predicted <ACH>  
F:97-581/Domain: chain A #status predicted <KR1>  
F:103-181/Domain: kringle homology <KR1>  
F:185-262/Domain: kringle homology <KR2>  
F:275-352/Domain: kringle homology <KR3>  
F:377-454/Domain: kringle homology <KR4>  
F:481-560/Domain: kringle homology <KR5>  
F:582-812/Domain: chain B #status predicted <BCH>  
F:582-803/Domain: trypsin homology <TRY>  
F:49-73-53-61-103-181-124-164-152-176-185-262-188-316-206-245-234-257-275-352-296-335-32  
bonds: #status predicted  
F:78-79/Cleavage site: Glu-Asn (stromelysin 1) #status predicted  
F:136-308/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:466-467/Cleavage site: Thr-Val (stromelysin 1) #status predicted  
F:581-582/Cleavage site: Arg-Val (plasminogen activator) #status experimental  
F:624-667-762/Active site: His, Asp, Ser #status predicted

Query Match 25.9%; Score 132; DB 1; Length 812;  
Best Local Similarity 33.7%; Pred. No. 1.6e-06;  
Matches 29; Conservative 13; Mismatches 36; Indels 8; Gaps 4;

QY 3 CYEGNGHYRGKASTDTMGRCLPWNSTVLOQTYVHAHRSALQGLGKHKYCNPNRR 62  
Db 377 CYQSDGQSYRGTSSTITTKKQCSW--AAMPFPHRSKTPENFPDAGL-EMNYCNPDPGX 433  
QY 63 RPWCYVQVGLKPLV--QECVHDCAD 86  
Db 434 GPWCYT---TPFSRWNYCNLKRSE 456

RESULT 48  
S33879  
plasmin precursor - lamprey (fragments)  
N:Alternate names: plasminogen  
C:Species: Petromyzonidae gen. sp. (lamprey)  
C>Date: 02-Dec-1993 #sequence\_revision 01-Sep-1995 #text\_change 07-Nov-1997  
C:Accession: S33879  
R:Affolter, M.; Schaller, J.; Rickli, E.E.  
A:Title: Isolation, characterization and partial amino acid sequence of lamprey plasmin  
A:Reference number: S33879  
A:Accession: S33879  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-15;16-34;35-44;45-59;60-76;77-111;111-138;139-158;159-178;179-216;217-236;  
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology  
F:81-146/Domain: kringle homology <KR3>

Query Match 25.7%; Score 131; DB 2; Length 336;  
Best Local Similarity 32.9%; Pred. No. 9.2e-07;  
Matches 28; Conservative 5; Mismatches 36; Indels 16; Gaps 2;

QY 3 CYEGNGHYRGKASTDTMGRCLPWNSTVLOQTYVHAHRSALQGLGKHKYCNPNRR 62  
Db 81 CVKGTGEGYRTALTIVSGKACQAWASQT-----PGDVTSCQGLVSNYCNPNPGEK 131  
QY 63 RPWCYVQVGLKPLVQECVHDCADG 87  
Db 132 LPWCYT-----TEYCNVPSTCG 149

RESULT 49  
JCS061  
macrophage-stimulating protein 1 precursor - rat

C:Species: Rattus norvegicus (Norway rat)  
C>Date: 31-Jan-1997 #sequence\_revision 31-Jan-1997 #text\_change 16-Jun-2000  
C:Accession: JCS061  
R:Ohshiro, K.; Iwama, A.; Matsumoto, K.; Ezaki, T.; Sakamoto, O.; Hamaguchi, I.;  
Biochem. Biophys. Res. Commun. 227, 273-280, 1996  
A:Title: Molecular cloning of rat macrophage-stimulating protein and its invol  
A:Reference number: JCS061; MUID:97011126; PMID:8958136  
A:Accession: JCS061  
A:Molecule type: mRNA  
A:Residues: 1-716 <OHS>  
A:Cross-references: EMBL:X95096; NID:G1669718; PIDN:CAA64473.1; PID:G1669719  
C:Complex: disulfide-bonded heterodimer of chains derived from the same precu  
C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology  
C:Keywords: duplication; glycoprotein; growth factor; kringle  
F:1-31/Domain: signal sequence #status predicted <SIG>  
F:32-488-489-716/Product: macrophage-stimulating protein 1 #status predicted <  
F:32-488/Domain: macrophage-stimulating protein 1 alpha chain #status predicted  
F:110-186/Domain: kringle homology <KR11>  
F:131-268/Domain: kringle homology <KR12>  
F:292-370/Domain: kringle homology <KR13>  
F:379-457/Domain: kringle homology <KR14>  
F:489-716/Domain: macrophage-stimulating protein 1 beta chain #status predicte  
F:489-709/Domain: trypsin homology <TRY>  
F:72-305-620/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 25.7%; Score 131; DB 1; Length 716;  
Best Local Similarity 33.8%; Pred. No. 1.9e-06;  
Matches 27; Conservative 10; Mismatches 23; Indels 20; Gaps 4;

QY 3 CYEGNGHYRGKASTDTMGRCLPWNSTVLOQTYVHAHRSALQGLGKHKY 54  
Db 292 CFRGKGYRGYTTNTTSAGVPCQRWDA----QNPQHRRFVPEKYACKDL-----RENF 340  
QY 55 CRPNPNRRPWCYV-QVGLK 73  
Db 341 CRPNPDGSEAPWCTTSRPGLR 360

RESULT 50  
IS1283  
hepatocyte growth factor precursor - clawed frog  
N:Alternate names: hepatietin A; scatter factor  
C:Species: Xenopus sp. (clawed frog)  
C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 18-Jun-1999  
C:Accession: IS1283  
R:Nakamura, H.; Tashiro, K.; Nakamura, T.; Shiohawa, K.  
Mech. Dev. 49, 123-131, 1995  
A:Title: Molecular cloning of Xenopus HGF cDNA and its expression studies in X  
A:Reference number: IS1283; MUID:95267650; PMID:7748783  
A:Accession: IS1283  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-710 <NAK>  
A:Cross-references: GB:S77422; NID:998932; PIDN:AAB34354.1; PID:998933  
A:Note: the authors' translation for residue 458 (Thr) is inconsistent with th  
C:Complex: disulfide-bonded heterodimer of chains derived from the same precu  
C:Function:  
A:Description: stimulates mitosis of hepatocytes and other cells  
A:Note: does not have proteinase activity  
C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology  
C:Keywords: duplication; glycoprotein; growth factor; heterodimer; kringle  
F:42-477-478-709/Product: hepatocyte growth factor #status predicted <MAT>  
F:42-477/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>  
F:115-193/Domain: kringle homology <KR1>  
F:198-275/Domain: kringle homology <KR2>  
F:289-367/Domain: kringle homology <KR3>  
F:375-453/Domain: kringle homology <KR4>  
F:478-709/Domain: hepatocyte growth factor beta chain #status predicted <BCH>  
F:478-700/Domain: trypsin homology <TRY>  
F:52-128-281-322-379-550-637-666/Binding site: carbohydrate (Asn) (covalent) #  
F:470-588/Disulfide bonds: #status predicted

Query Match 25.0%; Score 127.5; DB 1; Length 710;

Best Local Similarity 30.1%; Pred. No. 4.5e-06;  
Matches 31; Conservative 10; Mismatches 35; Indels 27; Gaps 5;  
QY 1 KTCVEGNHGFYRGKASTDTMGRPCLPWSATVLOQTYHAHR-----SDALQLGLGKH 52  
Db 287 KDCMKQGEYRGVSVITYNGIQCRWDS-----QFHLNFTPEYKCKDL-----SE 335  
QY 53 NYCRRNPDRRRPWCY-----VQVGLKPLVQECMV---HDCADG 87  
Db 336 NYCRRNPDSGSESPWCFPTTDPNIRIGHCSQIKKQASNQCECYG 378

Search completed: May 25, 2004, 14:58:28  
Job time : 7.09027 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2004, 14:43:40 ; Search time 3.39351 Seconds  
(without alignments)  
1350.274 Million cell updates/sec

Title: US-09-880-503-1  
Perfect score: 510  
Sequence: 1 KTCVEGNHGYRGKASTDTM.....QVGLKPLVQCMVHDCADGK 88

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	510	100.0	431	UROK_HUMAN	P00749 homo sapien
2	475	93.1	433	UROK_FAPCY	P16227 papio cynoc
3	420	82.4	442	UROK_PIG	P04185 sus scrofa
4	406	79.6	432	UROK_RAT	P29598 rattus norv
5	396	77.6	433	UROK_BOVIN	O05589 bos taurus
6	387	75.9	433	UROK_MOUSE	P05869 mus musculu
7	241	47.3	431	URTE_DESRO	P98121 desmodus ro
8	241	47.3	477	URTE_DESRO	P15638 desmodus ro
9	226	44.3	562	TPA_HUMAN	P00750 homo sapien
10	225	44.1	559	TPA_RAT	P19637 rattus norv
11	220	43.1	394	URTG_DESRO	P49150 desmodus ro
12	219	42.9	559	TPA_MOUSE	P11214 mus musculu
13	213	41.8	477	URTI_DESRO	P98119 desmodus ro
14	209	41.0	566	TPA_BOVIN	Q28198 bos taurus
15	198.5	38.9	434	UROK_CHICK	P15120 gallus gall
16	193	37.8	655	HGFA_HUMAN	Q04756 homo sapien
17	188	36.9	653	HGFA_MOUSE	Q91098 mus musculu
18	180.5	35.4	603	FA12_CAVPO	Q04962 cavia porce
19	170.5	33.4	615	FA12_HUMAN	P00748 homo sapien
20	161	31.6	1420	APOA_MACMU	P14417 macaca mula
21	159	31.2	473	KRM1_MOUSE	Q99n43 mus musculu
22	159	31.2	473	KRM1_RAT	Q924s4 rattus norv
23	158	31.0	452	KRM1_XENLA	Q90y90 xenopus lae
24	156	30.6	475	KRM1_HUMAN	Q96m18 homo sapien
25	156	30.6	4548	APOA_HUMAN	P08519 homo sapien
26	155.5	30.5	593	FA12_BOVIN	P98140 bos taurus
27	149.5	29.3	169	PLMN_RAT	Q01177 rattus norv
28	149	29.2	812	PLMN_BOVIN	P08868 bos taurus
29	148.5	29.1	462	PLMN_HUMAN	Q85c40 homo sapien
30	148.5	29.1	810	PLMN_ERIEU	Q29485 erinaceus e
31	148	29.0	790	PLMN_PIG	P06867 sus scrofa
32	147.5	28.9	943	ROR2_HUMAN	Q01974 homo sapien
33	147.5	28.9	944	ROR2_MOUSE	Q92138 mus musculu

ALIGNMENTS

RESULT 1		UROK_HUMAN		UROK_HUMAN		STANDARD;		PRT;		431 AA.	
ID	UROK_HUMAN	Q15944;	Q16618;	Q969W6;	AC	P00749;	Q15944;	Q16618;	Q969W6;	DT	21-JUL-1986 (Rel. 01, Created)
DT	21-JUL-1986 (Rel. 01, Created)	DT	20-MAR-1987 (Rel. 04, Last sequence update)	DT	10-OCT-2003 (Rel. 42, Last annotation update)	DE	Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)	DE	(U-plasminogen activator)	GN	PLAU.
OS	Homo sapiens (Human)	OS	Homo sapiens (Human)	OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	OX	NCBI_TaxID=9606;	OX	NCBI_TaxID=9606;	RP	SEQUENCE FROM N.A.
RA	Riccio A., Grimaldi G., Verde P., Sebastiao G., Boast S., Blasi F.;	RA	Riccio A., Grimaldi G., Verde P., Sebastiao G., Boast S., Blasi F.;	RT	"The human urokinase-plasminogen activator gene and its promoter.";	RL	Nucleic Acids Res. 13:2759-2771(1985).	RN	[1]	RN	Q924s4 rattus norv
RT	"Cloning and expression of the gene for pro-urokinase in Escherichia coli.";	RT	colli.";	RT	colli.";	RT	colli.";	RT	colli.";	RT	colli.";



RL Biotechnology 3:923-929(1985).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86056954; PubMed=2415429;  
RA Nagai M., Hiramatsu R., Kaneda T., Hayasuke N., Atimura H.,  
RA Nishida M., Suyama T.;  
RT "Molecular cloning of cDNA coding for human preprourokinase.";  
RL Gene 36:183-188(1985).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85203359; PubMed=3888571;  
RA Jacobs P., Cravador A., Loriau R., Brockly F., Colau B., Chuchana P.,  
RA van Elsen A., Herzog A., Bollen A.;  
RT "Molecular cloning, sequencing, and expression in *Escherichia coli* of  
ET human preprourokinase cDNA.";  
RL DNA 4:139-146(1985).  
RN [5]  
RP SEQUENCE FROM N.A.  
RA Rieder M.J., Carrington D.F., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,  
RA Nickerson D.A.;  
RN Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RX TISSUE=lung;  
RC MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loguailano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Pahay J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
ET human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [7]  
RP SEQUENCE OF 66-431 FROM N.A.  
RX MEDLINE=84272706; PubMed=6589620;  
RA Verde P., Stoppelli M.P., Galeffi P., di Nocera P., Blasi F.;  
RT "Identification and primary sequence of an unspliced human urokinase  
ET poly(A)+ RNA.";  
RL Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).  
RN [8]  
RP SEQUENCE OF 21-177.  
RX MEDLINE=83055084; PubMed=6754569;  
RA Gunzler W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E.,  
RA Flohe L.;  
RT "The primary structure of high molecular mass urokinase from human  
ET urine. The complete amino acid sequence of the A chain.";  
RL Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).  
RN [9]  
RP SEQUENCE OF 156-176 AND 179-224.  
RX MEDLINE=83003608; PubMed=6749491;  
RA Schaller J., Nick H., Rickli E.E., Gillesen D., Lergier W.,  
RA Studer R.O.;  
RT "Human low-molecular-weight urinary urokinase. Partial  
ET characterization and preliminary sequence data of the two polypeptide  
RT chains.";  
RL Eur. J. Biochem. 125:251-257(1982).  
RN [10]  
RP SEQUENCE OF 158-410.  
RX MEDLINE=83055099; PubMed=6754572;  
RA Steffens G.J., Gunzler W.A., Otting F., Frankus E., Flohe L.;  
RA

RT "The complete amino acid sequence of low molecular mass urokinase  
ET from human urine.";  
RL Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).  
RN [11]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).  
RX MEDLINE=96000858; PubMed=8591045;  
RA Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,  
RA Dobson C.M., Stuart D.I., Jones E.Y.;  
RT "The crystal structure of the catalytic domain of human  
ET urokinase-type plasminogen activator.";  
RL Structure 3:681-691(1995).  
RN [12]  
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.  
RX MEDLINE=20366327; PubMed=10805774;  
RA Sperl S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G.,  
RA Bode W., Magdolen V., Huber R., Moroder L.;  
RT "(4-aminomethyl)phenylguanidine derivatives as nonpeptidic highly  
ET selective inhibitors of human urokinase.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).  
RN [13]  
RP STRUCTURE BY NMR.  
RX MEDLINE=89127526; PubMed=2536903;  
RA Oswald R.E., Bogusky M.J., Bamberger M., Smith R.A.G., Dobson C.M.;  
RT "Dynamics of the multidomain fibrinolytic protein urokinase from two-  
ET dimensional NMR.";  
RL Nature 337:579-582(1989).  
RN [14]  
RP STRUCTURE BY NMR OF 67-155.  
RX MEDLINE=93003110; PubMed=1327118;  
RA Li X., Smith R.A.G., Dobson C.M.;  
RT "Sequential 1H NMR assignments and secondary structure of the kringle  
ET domain from urokinase.";  
RL Biochemistry 31:9562-9571(1992).  
RN [15]  
RP STRUCTURE BY NMR OF 67-155.  
RX MEDLINE=94149701; PubMed=8107091;  
RA Li X., Bokman A.M., Ilinas M., Smith R.A.G., Dobson C.M.;  
RT "Solution structure of the kringle domain from urokinase-type  
ET plasminogen activator.";  
RL J. Mol. Biol. 235:1548-1559(1994).  
RN [16]  
RP VARIANT LEU-141.  
RX MEDLINE=96186279; PubMed=8652631;  
RA Yoshimoto M., Ushiyama Y., Sakai M., Tamaki S., Hara H., Takahashi K.,  
RA Sawasaki Y., Hanada K.;  
RT "Characterization of single chain urokinase-type plasminogen  
ET activator with a novel amino-acid substitution in the kringle  
ET structure.";  
RL Biochim. Biophys. Acta 1293:83-89(1996).  
RN [17]  
RP VARIANT LEU-141.  
RX MEDLINE=97218551; PubMed=9065988;  
RA Conne B., Berczy M., Belin D.;  
RT "Detection of polymorphisms in the human urokinase-type plasminogen  
ET activator gene.";  
RL Thromb. Haemost. 77:434-435(1997).  
RN [18]  
RP ERRATUM.  
RA Conne B., Berczy M., Belin D.;  
RL Thromb. Haemost. 78:973-973(1997).  
RN [19]  
RP VARIANT LEU-141.  
RX MEDLINE=97337920; PubMed=9194591;  
RA Turkmen B., Schmitt M., Schmalfeldt B., Trommler P., Heil W.,  
RA Creutzburg S., Graeff H., Magdolen V.;  
RT "Mutational analysis of the genes encoding urokinase-type plasminogen  
ET activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer.";  
RL Electrophoresis 18:686-689(1997).  
RN [20]  
RP FUNCTION: Potent plasminogen activator and is clinically used for  
ET therapy of thrombolytic disorders.  
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
CC plasminogen to form plasmin.  
CC -1- SUBUNIT: Found in high and low molecular mass forms. Each consists

CC of two chains, A and B. The high molecular mass form contains a  
 CC long chain A. Cleavage occurs after residue 155 in the low  
 CC molecular mass form to yield a short A1 chain.  
 CC -!- PHARMACEUTICAL: Available under the name Abbokinase (Abbott). Used  
 CC in Pulmonary Embolism (PE) to initiate fibrinolysis.  
 CC -!- SIMILARITY: Belongs to peptidase family S1.  
 CC -!- SIMILARITY: Contains 1 EGF-like domain.  
 CC -!- SIMILARITY: Contains 1 kringle domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X51935; CAA36200.1; -  
 CC PIR; S14687; UKBAY.  
 CC HSP; P00749; ILMW.  
 CC MEROPS; S01.231; -  
 CC InterPro; IPR009003; Cys Ser trypsin.  
 CC InterPro; IPR006209; EGF-like.  
 CC InterPro; IPR006210; IEGF.  
 CC InterPro; IPR000001; Kringle.  
 CC InterPro; IPR008293; Pept\_S1A\_uPA.  
 CC InterPro; IPR001254; Peptidase S1.  
 CC InterPro; IPR001314; Peptidase\_S1A.  
 CC Pfam; PF00051; kringle; 1.  
 CC Pfam; PF00089; trypsin; 1.  
 CC PIRSF; PIRSF001144; Urk Plasm act; 1.  
 CC PRINTS; PR00722; CHYMOTRYPSIN.  
 CC PRINTS; PR00018; Kringle.  
 CC ProDom; PD000395; Kringle; 1.  
 CC SMART; SM00181; EGF; 1.  
 CC SMART; SM00130; KR; 1.  
 CC SMART; SM00020; Tryp\_SPC; 1.  
 CC PROSITE; PS00022; EGF\_1; 1.  
 CC PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
 CC PROSITE; PS00026; EGF\_3; 1.  
 CC PROSITE; PS00021; KRINGLE\_1; 1.  
 CC PROSITE; PS00070; KRINGLE\_2; 1.  
 CC PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 CC PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 CC PROSITE; PS00135; TRYPSIN\_SER; 1.  
 CC Plasmicogen activation; Hydrolase; Serine protease; Glycoprotein;  
 CC Kringle; EGF-like domain; Zymogen; Signal.  
 CC SIGNAL 1 20  
 CC CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.  
 CC CHAIN 21 176 CHAIN A (BY SIMILARITY).  
 CC CHAIN 155 176 SHORT A CHAIN (A1) (BY SIMILARITY).  
 CC CHAIN 178 433 CHAIN B (BY SIMILARITY).  
 CC DOMAIN 26 62 EGF-LIKE.  
 CC DOMAIN 69 150 KRINGLE.  
 CC DOMAIN 151 177 CONNECTING PEPTIDE.  
 CC DOMAIN 178 433 SERINE PROTEASE.  
 CC DISULFID 30 38 BY SIMILARITY.  
 CC DISULFID 32 50 BY SIMILARITY.  
 CC DISULFID 52 61 BY SIMILARITY.  
 CC DISULFID 167 298 INTERCHAIN (BY SIMILARITY).  
 CC DISULFID 208 224 BY SIMILARITY.  
 CC DISULFID 216 287 BY SIMILARITY.  
 CC DISULFID 315 384 BY SIMILARITY.  
 CC DISULFID 347 363 BY SIMILARITY.  
 CC DISULFID 374 402 BY SIMILARITY.  
 CC ACT\_SITE 223 CHARGE RELAY SYSTEM.  
 CC ACT\_SITE 274 CHARGE RELAY SYSTEM.  
 CC ACT\_SITE 378 CHARGE RELAY SYSTEM.  
 CC CARBOHYD 324 324 N-LINKED (GLCNAC...) (BY SIMILARITY).  
 CC SQ SEQUENCE 433 AA; 48595 MW; 816D22DFEDDC8792 CRC64;

Query Match 100.0%; Score 510; DB 1; Length 431;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-51;  
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KTCVEGNHFGYRGKASTDTMGRCLPWSATVLTQOYTHAHSALQGLGKHNYCRNPDN 60  
 DB 68 KTCVEGNHFGYRGKASTDTMGRCLPWSATVLTQOYTHAHSALQGLGKHNYCRNPDN 127.  
 QY 61 RRPWCYVQGLKPLVQECMVHDCADGK 88  
 DB 128 RRPWCYVQGLKPLVQECMVHDCADGK 155

RESULT 2  
 UROK PAPCY STANDARD; PRT; 433 AA.  
 AC P16227;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)  
 DE (U-plasminogen activator).  
 GN P1AU.  
 OS Papio cynocephalus (Yellow baboon).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecoidea; Papio.  
 OC NCBI\_TaxID=9556;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Thoracic aorta;  
 RX MEDLINE=90287734; PubMed=2113276;  
 RA Au Y. P. T., Wang T. W., Clowes A. W.;  
 RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type  
 RT plasminogen activator";  
 RL Nucleic Acids Res. 18:3411-3411(1990).  
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
 CC plasminogen to form plasmin.  
 CC -!- SUBUNIT: Found in high and low molecular mass forms. Each consists  
 CC of two chains, A and B. The high molecular mass form contains a  
 CC long chain A. Cleavage occurs after residue 155 in the low  
 CC molecular mass form to yield a short A1 chain (By similarity).

Query Match 93.1%; Score 475; DB 1; Length 433;  
 Best Local Similarity 94.3%; Pred. No. 2.7e-47;  
 Matches 83; Conservative 1; Mismatches 4; Indels 0; Gaps 0  
 QY 1 KTCVEGNHFGYRGKASTDTMGRCLPWSATVLTQOYTHAHSALQGLGKHNYCRNPDN 60  
 DB 67 KTCVEGNHFGYRGKASTDTMGRCLPWSATVLTQOYTHAHSALQGLGKHNYCRNPDN 126

```

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 127 RRRPWCYVQVGLKQVQECMVHNCADGK 154

RESULT 3
UOK_PIG STANDARD; PRT; 442 AA.
AC P04185;
DT 20-MAR-1987 (Rel. 04, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)
DE (U-plasminogen activator).
GN PLA0.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
EX MEDLINE=85087954; PubMed=6096932;
RA Nagamine Y., Pearson D., Altus M.S., Reich E.;
RT "cDNA and gene nucleotide sequence of porcine plasminogen activator.";
RL Nucleic Acids Res. 12:9525-9541(1984).
RN [2]
RP REVISION TO 241.
RA Nagamine Y.;
RL Submitted (DEC-1986) to the PIR data bank.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL; X01648; CAA25806.1; -
DR EMBL; X02724; CAA26511.1; -
DR PIR; A00932; UKPO.
DR HSSP; P00749; 1KDU.
DR MEROPS; S01.231; -.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR008293; Peptidase S1.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PIRSF; PIRSF001144; Urk Plasmin act; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS50026; EGF_3; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPsin DOM; 1.
DR PROSITE; PS00134; TRYPsin SER; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;

KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20
FT CHAIN 21 442
FT CHAIN 21 198
FT CHAIN 21 198
FT CHAIN 190 442
FT DOMAIN 29 65
FT DOMAIN 72 153
FT DOMAIN 154 189
FT DOMAIN 190 442
FT CARBOHYD 152 152
FT DISULFID 33 41
FT DISULFID 35 53
FT DISULFID 55 64
FT DISULFID 179 310
FT DISULFID 220 236
FT DISULFID 228 239
FT DISULFID 324 333
FT DISULFID 356 372
FT DISULFID 383 411
FT ACT_SITE 235 235
FT ACT_SITE 286 286
FT ACT_SITE 387 387
FT CONFLICT 241 241
FT CONFLICT 242 242
FT CONFLICT 288 288
SQ SEQUENCE 442 AA; 49116 MW; EE32FCFEF501321EE CRC64;

Query Match 82.4%; Score 420; DB 1; Length 442;
Best Local Similarity 81.8%; Pred. No. 6e-41; Gaps 0.
Matches 72; Conservative 8; Mismatches 8; Indels 0; Gaps 0.

QY 1 KTCYEGNGHYRGKASTDTMGRCPLPMSATVLOQTYHAHSDALQLGLGKHNYCRNPDN 60
Db 70 QTCFEGNGHSYRGKANTTGTGRPCPLPMSATVLLNTYHAHSDALQLGLGKHNYCRNPDN 129
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 130 QRRPWCYVQVGLKPLVQECMVHNCADGK 157

RESULT 4
UOK_RAT STANDARD; PRT; 432 AA.
AC P29598;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)
DE (U-plasminogen activator).
GN PLA0.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pi-scher 344;
RC MEDLINE=92233409; PubMed=1568219;
RA Henderson B.R., Tansey W.P., Phillips S.M., Ramshaw I.A.,
RA Keifford R.F.;
RT "Transcriptional and posttranscriptional activation of urokinase
RT plasminogen activator gene expression in metastatic tumor cells.";
RL Cancer Res. 52:2489-2496(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Rabbani S.A.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBSJ databases.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: Found in high and low molecular mass forms. Each consists
CC of two chains, A and B. The high molecular mass form contains a
CC long chain A. Cleavage occurs after residue 156 in the low

```

molecular mass form to yield a short A1 chain (By similarity).  
 -!- SIMILARITY: Belongs to peptidase family S1.  
 -!- SIMILARITY: Contains 1 EGF-like domain.  
 -!- SIMILARITY: Contains 1 kringle domain.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; X63434; CAA45028.1; -;  
 EMBL; X65651; CAA46601.1; -;  
 PIR; S24604; S18932.  
 HSSP; P00749; 1KDU.  
 MEROPS; S01.231; -;  
 InterPro; IPR009003; Cys Ser trypsin.  
 InterPro; IPR006209; EGF-like.  
 InterPro; IPR006210; IEGF.  
 InterPro; IPR000001; Kringle.  
 InterPro; IPR008293; Pept S1A uPA.  
 InterPro; IPR001254; Peptidase S1.  
 InterPro; IPR001314; Peptidase\_S1A.  
 Pfam; PF00051; kringle; 1.  
 Pfam; PF00089; trypsin; 1.  
 PIRSF; PIRSF001144; Urk\_plasm\_act; 1.  
 PRINTS; PR00722; CHYMOTRYPSIN.  
 PRINTS; PR00018; KRINGLE.  
 ProDom; PD000395; Kringle; 1.  
 SMART; SM00181; EGF; 1.  
 SMART; SM00130; KR; 1.  
 SMART; SM00020; Tryp\_Spc; 1.  
 PROSITE; PS00022; EGF\_1;  
 PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
 PROSITE; PS00026; EGF\_3; 1.  
 PROSITE; PS00021; KRINGLE\_1; 1.  
 PROSITE; PS00070; KRINGLE\_2; 1.  
 PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 PROSITE; PS00134; TRYPSIN\_HIS; FALSE\_NEG.  
 PROSITE; PS00135; TRYPSIN\_SER; 1.  
 Kringle; EGF-like domain; Hydrolase; Serine protease; Glycoprotein;  
 Kringle; EGF-like domain; Zymogen; Signal.  
 SIGNAL 1 19 POTENTIAL.  
 CHAIN 20 432 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.  
 CHAIN A (BY SIMILARITY)  
 SHORT A CHAIN (A1) (BY SIMILARITY).  
 CHAIN B (BY SIMILARITY).  
 EGF-LIKE.  
 KRINGLE.  
 CONNECTING PEPTIDE.  
 SERINE PROTEASE.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 INTERCHAIN (BY SIMILARITY).  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 CHARGE RELAY SYSTEM.  
 CHARGE RELAY SYSTEM.  
 CHARGE RELAY SYSTEM.  
 E -> H (IN REF. 2).  
 E -> G (IN REF. 2).  
 D -> N (IN REF. 2).  
 SEQUENCE 432 AA; 47957 MW; 4EB1B96C716244C6 CRC64;

Query March 79.6%; Score 406; DB 1; Length 432;  
 Best Local Similarity 79.5%; Pred. No. 2.4e-39;  
 Matches 70; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

QY 1 KTCYEGNHFYRGKASTDTMGPCPLPMSATVLOQTYHAHRSDALQIGLGHNYCRNDP 60  
 DB 68 KTCYHGNGSGYKGNANTDKRPRCLAWNSPAVLOQTYHAHRSDALSGLGHNYCRNDP 127  
 QY 61 RRRPCWYVQVGLKPLVQECMVHDCADGK 88  
 DB 128 QRRPCWYVQVGLKPLVQECMVHDCADGK 155  
 RESULT 5  
 UROK\_BOVIN STANDARD; PRT; 433 AA.  
 ID UROK\_BOVIN STANDARD; PRT; 433 AA.  
 AC C05589; Q28209;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)  
 DE U-plasminogen activator).  
 GN FLAU.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OC NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Aortic endothelium;  
 RX MEDLINE=93216119; PubMed=8385052;  
 RA Kraetzschmar J., Haendler B., Kojima S., Rifkin D.B.,  
 RA Schleuning W.-D.;  
 RT "Bovine urokinase-type plasminogen activator and its receptor:  
 cloning and induction by retinoic acid.";  
 RL Gene 125:177-183(1993).  
 RN [2]  
 RP SEQUENCE OF 12-433 FROM N.A.  
 RC TISSUE=Kidney;  
 RA Ravn P., Berglund L., Petersen T.E.;  
 RT "Cloning and characterization of the bovine plasminogen activators uPA  
 and tPA.";  
 CC Int. Dairy J. 5:605-617(1995).  
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
 plasminogen to form plasmin.  
 CC -!- INDUCTION: By retinoic acid.  
 CC -!- SIMILARITY: Belongs to peptidase family S1.  
 CC -!- SIMILARITY: Contains 1 EGF-like domain.  
 CC -!- SIMILARITY: Contains 1 kringle domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; I03546; AAA51419.1; -;  
 EMBL; X85801; CAA59796.1; -;  
 PIR; JNC0560; JNC0560.  
 HSSP; P00749; 1LMW.  
 MEROPS; S01.231; -;  
 InterPro; IPR009003; Cys Ser trypsin.  
 InterPro; IPR006209; EGF-like.  
 InterPro; IPR000001; Kringle.  
 InterPro; IPR008293; Pept S1A uPA.  
 InterPro; IPR001254; Peptidase S1.  
 InterPro; IPR001314; Peptidase\_S1A.  
 Pfam; PF00051; kringle; 1.  
 Pfam; PF00089; trypsin; 1.  
 PIRSF; PIRSF001144; Urk\_plasm\_act; 1.  
 PRINTS; PR00722; CHYMOTRYPSIN.  
 PRINTS; PR00018; KRINGLE.  
 ProDom; PD000395; Kringle; 1.

```
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRYP_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS00026; EGF_3; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYP_SIN_DOM; 1.
DR PROSITE; PS00134; TRYP_SIN_HIS; 1.
DR PROSITE; PS00135; TRYP_SIN_SER; 1.
KW Plasminogen activator; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Zymogen.
FT SIGNAL 1 20
FT CHAIN 21 433
FT CHAIN 21 179
FT CHAIN 181 433
FT DOMAIN 29 65
FT DOMAIN 72 153
FT DOMAIN 154 180
FT DOMAIN 181 433
FT DISULFID 33 41
FT DISULFID 35 53
FT DISULFID 55 64
FT DISULFID 170 301
FT DISULFID 211 227
FT DISULFID 219 290
FT DISULFID 315 384
FT DISULFID 347 363
FT DISULFID 374 402
FT ACT SITE 226 226
FT ACT SITE 277 277
FT ACT SITE 378 378
FT CONFLICT 189 189
FT SEQUENCE 433 AA; 48730 MW; 4DE1B8D4DA47027A CRC64;

Query Match 77.6%; Score 396; DB 1; Length 433;
Best Local Similarity 76.1%; Pred. No. 3.4e-38;
Matches 67; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCLPWNSATVLQOTYHRSALQGLGKHYCRNPDN 60
DB 70 KTCYQNGHSTYRGKANDLGRCLPWNSATVLQOTYHRSALQGLGKHYCRNPDN 129
QY 61 RRRPWCYVQGLKPLVQECMVHDCADGK 88
DB 130 QRRPWCYVQGLKQFVQCMVQDCSVGK 157

RESULT 6
UROK_MOUSE STANDARD; PRT; 433 AA.
AC P06869;
DT 01-JAN-1998 (Rel. 06, Created)
DT 01-JAN-1998 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN FLAU.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85179474; PubMed=2985383;
RA Bein D., Vassalli J.-D., Combepine C., Godeau F., Nagamine Y.,
RA Reich E., Kocher H.P., Duvoisin R.M.;
RT "Cloning, nucleotide sequencing and expression of cDNAs encoding
RT mouse urokinase-type plasminogen activator.";
RL Eur. J. Biochem. 148:225-232(1985).
[2]
RN SEQUENCE FROM N.A.
RP MEDLINE=88163489; PubMed=2831940;
RX
```

```
RA Degen S.J.F., Heckel J.L., Reich E., Degen J.L.;
RT "The murine urokinase-type plasminogen activator gene.";
RL Biochemistry 26:8270-8279(1987).
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: Found in high and low molecular mass forms. Each consists
CC of two chains, A and B. The high molecular mass form contains a
CC long chain A. Cleavage occurs after residue 156 in the low
CC molecular mass form to yield a short A1 chain (By similarity).
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercia
CC entities requires a license agreement (See http://www.isb-sib.ch/announce.
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; X02389; CAA26231.1; -.
CC EMBL; M17922; AAA40539.1; -.
CC PIR; A29420; UKMS.
CC HSSP; P00749; 1KDU.
CC MEROPS; S01.231; -.
CC MGD; MGI:97611; Flau.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR008293; Pept_S1A_UPA.
CC InterPro; IPR001454; Peptidase_S1.
CC InterPro; IPR001314; Peptidase_S1A.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC PIRSF; PIRSF001144; Urk_plasm_act; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00181; EGF; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00020; TRYP_SPC; 1.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; FALSE_NEG.
CC PROSITE; PS00026; EGF_3; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS00070; KRINGLE_2; 1.
CC PROSITE; PS00240; TRYP_SIN_DOM; 1.
CC PROSITE; PS00134; TRYP_SIN_HIS; FALSE_NEG.
CC PROSITE; PS00135; TRYP_SIN_SER; 1.
KW Plasminogen activator; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20
FT CHAIN 21 433
FT CHAIN 21 178
FT CHAIN 157 178
FT CHAIN 180 433
FT CHAIN 180 433
FT DOMAIN 28 64
FT DOMAIN 71 152
FT DOMAIN 153 179
FT DOMAIN 180 433
FT DISULFID 32 40
FT DISULFID 34 52
FT DISULFID 54 63
FT DISULFID 169 301
FT DISULFID 211 227
FT DISULFID 219 290
FT DISULFID 315 384
FT DISULFID 347 363
FT ACT SITE 226 226
FT ACT_SITE 277 277
```

```

FT ACT SITE 378 378 CHARGE RELAY SYSTEM.
SQ SEQUENCE 433 AA; 48268 MW; A39C35F6250443F9 CRC64;

Query Match 75.9%; Score 387; DB 1; Length 433;
Best Local Similarity 72.7%; Pred. No. 3.8e-37;
Matches 64; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNYCRPNDR 60
Db KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNYCRPNDR 128

QY 51 RRRPWCYVQGLKPLVQECWHDCA 88
Db RRRPWCYVQGLKPLVQECWHDCA 156

RESULT 7
URTB_DESRO STANDARD; PRT; 431 AA.
AC P98121;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator beta precursor (EC 3.4.21.68) (DSPA
DE beta).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=9339059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC -!- FUNCTION: Probably essential to support the feeding habits of this
CC exclusively haematophagous animal. Probable potent thrombolytic
CC agent.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg|-Val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: Monomer.
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M61989; AAA31594.1; -.
CC PIR; J50599; J50599.
CC HSRF; P98119; IAS1.
CC MEKOPS; S01.239; -.
CC InterPro; IPR009003; Cys Ser trypsin.
CC InterPro; IPR006209; EGF-like.
CC InterPro; IPR006210; IEGF.

```

```

DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; trypt; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Multigene family.
FT SIGNAL 1 36
FT CHAIN 37 431
FT DOMAIN 37 75
FT DOMAIN 82 163
FT DOMAIN 179 431
FT ACT_SITE 226 226
FT ACT_SITE 275 275
FT ACT_SITE 382 382
FT DISULFID 41 52
FT DISULFID 46 63
FT DISULFID 65 74
FT DISULFID 82 163
FT DISULFID 103 145
FT DISULFID 134 158
FT DISULFID 168 299
FT DISULFID 211 227
FT DISULFID 219 288
FT DISULFID 313 388
FT DISULFID 345 361
FT DISULFID 378 406
FT CARBOHYD 139 139
FT CARBOHYD 352 352
SQ SEQUENCE 431 AA; 48221 MW; 699B5B675B162CBF CRC64;

Query Match 47.3%; Score 241; DB 1; Length 431;
Best Local Similarity 50.0%; Pred. No. 2.6e-20;
Matches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0

QY 2 TCYEGNGHYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNYCRPNDR 61
Db TCYEGNGHYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNYCRPNDR 140

QY 62 RRPWCYVQGLKPLVQECWHDCA 85
Db RRPWCYVQGLKPLVQECWHDCA 164

RESULT 8
URTB_DESRO STANDARD; PRT; 477 AA.
ID UR2_DESRO
AC P15638;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator alpha 2 precursor (EC 3.4.21.68) (DSPA
DE alpha-2) (BAT-PA) (T-plasminogen activator).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;

```

RN SEQUENCE FROM N.A.  
RP TISSUE=Salivary gland;  
RC MEDLINE=92039036; PubMed=1937019;  
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,  
RA Alagon A., Donner P., Schleuning W.D.;  
RT "the plasminogen activator family from the salivary gland of the  
RT vampire bat Desmodus rotundus: cloning and expression.";  
RL Gene 105:229-237(1991).  
RN [2]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Salivary gland;  
EX MEDLINE=90036967; PubMed=2509450;  
RA Gardell S.J., Duong L.T., Diehl R.E., York J.D., Hare T.R.,  
RA Register R.B., Jacobs J.W., Dixon R.A.F., Friedman P.A.;  
RT "Isolation, characterization, and cDNA cloning of a vampire bat  
RT salivary plasminogen activator.";  
RL J. Biol. Chem. 264:17947-17952(1989).  
RN [3]  
RP CHARACTERIZATION.  
EX MEDLINE=9393059; PubMed=1309059;  
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,  
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,  
RA Donner P.;  
RT "plasminogen activators from the saliva of Desmodus rotundus (common  
RT vampire bat): unique fibrin specificity.";  
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).  
CC -!- FUNCTION: Probably essential to support the feeding habits of this  
CC exclusively haematophagous animal. Probable potent thrombolytic  
CC agent.  
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
CC plasminogen to form plasmin.  
CC -!- ENZYME REGULATION: Activity toward plasminogen is stimulated in  
CC the presence of fibrin I.  
CC -!- SUBUNIT: Monomer.  
CC -!- DOMAIN: The fibronectin type-I domain mediates binding to fibrin,  
CC and the kringle domain apparently mediates fibrin-induced  
CC stimulation of activity.  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
CC -!- SIMILARITY: Contains 1 EGF-like domain.  
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.  
CC -!- SIMILARITY: Contains 1 kringle domain.  
CC  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M63988; AAA31593.1; -;  
CC EMBL; J05082; AAA31596.1; -;  
CC PIR; A34369; A34369.  
CC PIR; J05098; J05098.  
CC HSP; P98119; 1A51.  
CC MEROPS; S01.232; -;  
CC InterPro; IPR009003; Cys\_Ser\_trypsin.  
CC InterPro; IPR006209; EGF-like.  
CC InterPro; IPR000083; Fibronctn1.  
CC InterPro; IPR006210; IEGR.  
CC InterPro; IPR000001; Kringle.  
CC InterPro; IPR001254; Peptidase\_S1.  
CC InterPro; IPR001314; Peptidase\_S1a.  
CC Pfam; PF00008; EGF; 1.  
CC Pfam; PF00039; fn1; 1.  
CC Pfam; PF00051; kringle; 1.  
CC Pfam; PF00089; trypsin; 1.  
CC PRINTS; PR00722; CHYMOTRYPSIN.  
CC PRINTS; PR00018; KRINGLE.  
CC ProDom; PD000395; Kringle; 1.  
CC SMART; SM00181; EGF; 1.  
CC SMART; SM00058; FN1; 1.

DR SMART; SM00130; KR; 1.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS01186; EGF\_2; 1.  
DR PROSITE; PS00026; EGF\_3; 1.  
DR PROSITE; PS01253; FIBRONECTIN\_1; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS00070; KRINGLE\_2; 1.  
DR PROSITE; PS00340; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW plasminogen activator; Hydrolase; Serine protease; Glycoprotein;  
KW kringle; EGF-like domain; Signal; Multigene family.  
FT SIGNAL 1 36  
FT CHAIN 37 477  
FT DOMAIN 40 82  
FT DOMAIN 83 121  
FT DOMAIN 128 209  
FT DOMAIN 225 477  
FT ACT\_SITE 272 272  
FT ACT\_SITE 321 321  
FT ACT\_SITE 428 428  
FT DISULFID 42 72  
FT DISULFID 70 79  
FT DISULFID 87 98  
FT DISULFID 92 109  
FT DISULFID 111 120  
FT DISULFID 128 209  
FT DISULFID 149 191  
FT DISULFID 180 204  
FT DISULFID 214 345  
FT DISULFID 257 273  
FT DISULFID 265 334  
FT DISULFID 359 434  
FT DISULFID 391 407  
FT DISULFID 424 452  
FT CARBOHYD 185 185  
FT CARBOHYD 398 398  
FT CONFLICT 403 403  
FT CONFLICT 417 417  
FT CONFLICT 435 435  
FT CONFLICT M -> R (IN REF. 2).  
SQ SEQUENCE 477 AA; 53719 MW; 17486555C0E5077C CRC64;  
Query Match 47.3%; Score 241; DB 1; Length 477;  
Best Local Similarity 50.0%; Pred. No. 2.9e-20;  
Matches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0.  
Qy 2 TCYEGNGHFVRGKASTDTWGRPCLPWSATVLQCTVHAHRSDALQGLGHNCRPNDR 61  
Db 127 TCYDQGVYRGTWSTSESGAQCNVNSLLRTYNGRSDAITLGLGHNCRPNDR 186  
Qy 62 RRFWCYVQVLKPLVQECMVHDC 85  
Db 187 SKPCYVIRAKSFLEFCSPVCS 210  
RESULT 9  
ID TPA\_HUMAN  
AC P00750; Q15103;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Tissue-type plasminogen activator (EC 3.4.21.68) (tpa)  
DE (t-PA) (t-plasminogen activator) (Alteplase) (Retenase).  
GN PLAT.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Melanoma;

RX MEDLINE=83115262; PubMed=6337343;  
RA Pernica D., Holmes W.E., Kohr W.J., Harkins R.N., Vehar G.A.,  
RA Ward C.A., Bennett W.F., Yelverton E., Seeburg P.H., Heyneker H.L.,  
RA Goeddel D.V., Collen D.;  
RT "Cloning and expression of human tissue-type plasminogen activator  
RT cDNA in *E. coli*.";  
RL Nature 301:214-221(1983).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fetal lung;  
RX MEDLINE=88262579; PubMed=3133640;  
RA Sasaki H., Saito Y., Hayashi M., Otsuka K., Niwa M.;  
RT "Nucleotide sequence of the tissue-type plasminogen activator cDNA  
RT from human fetal lung cells.";  
RL Nucleic Acids Res. 16:5695-5695(1988).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86054470; PubMed=2824147;  
RA Reddy V.B., Garramone A.J., Sasak H., Wei C.M., Watkins P., Galli J.,  
RA Hsiung N.;  
RT "Expression of human uterine tissue-type plasminogen activator in  
RT mouse cells using BPV vectors.";  
RL DNA 6:461-472(1987).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86196143; PubMed=3009482;  
RA Friezen Degen S.J., Rajput B., Reich E.;  
RT "The human tissue plasminogen activator gene.";  
RL J. Biol. Chem. 261:6972-6985(1986).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=84298137; PubMed=6089198;  
RA NY T., Elgh P., Lund B.;  
RT "The structure of the human tissue-type plasminogen activator gene:  
RT correlation of intron and exon structures to functional and  
RT structural domains.";  
RL Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359(1984).  
RN [6]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86284200; PubMed=3090401;  
RA Harris T.J., Patel T., Marston F.A., Little S., Emtage J.S.,  
RA Odenakker G., Volckaert G., Ronbouts W., Billiau A., Somer P.;  
RT "Cloning of cDNA coding for human tissue-type plasminogen activator  
RT and its expression in *Escherichia coli*.";  
RL Mol. Biol. Med. 3:279-292(1986).  
RN [7]  
RP SEQUENCE FROM N.A. (ISOFORM SHORT).  
RC TISSUE=Umbilical vein;  
RX MEDLINE=90192129; PubMed=2107528;  
RA Siebert P.D., Fong K.;  
RT "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from  
RT human endothelial cells.";  
RL Nucleic Acids Res. 18:1086-1086(1990).  
RN [8]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Raha S.S., Iaquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettawan M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grumwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [9]  
RP SEQUENCE OF 212-361 FROM N.A.  
RX MEDLINE=83169656; PubMed=6572897;  
RA Edlund T., Ny T., Raanby M., Heden L.-O., Palm G., Holmgren E.,  
RA Josephson S.;  
RT "Isolation of cDNA sequences coding for a part of human tissue  
RT plasminogen activator.";  
RL Proc. Natl. Acad. Sci. U.S.A. 80:349-352(1983).  
RN [10]  
RP SEQUENCE OF 1-36 FROM N.A.  
RX MEDLINE=85289338; PubMed=3161893;  
RA Fisher R., Waller E.K., Grossi G., Thompson D., Tizard R.,  
RA Schleuning W.-D.;  
RT "Isolation and characterization of the human tissue-type plasminogen  
RT activator structural gene including its 5' flanking region.";  
RL J. Biol. Chem. 260:11223-11230(1985).  
RN [11]  
RP SEQUENCE OF 31-562 FROM N.A.  
RX MEDLINE=91291340; PubMed=1368681;  
RA Itagaki Y., Yasuda H., Morinaga T., Mitsuoka S., Higashio K.;  
RT "Purification and characterization of tissue plasminogen activator  
RT secreted by human embryonic lung diploid fibroblasts, IMR-90 cells.";  
RL Agric. Biol. Chem. 55:1225-1232(1991).  
RN [12]  
RP SEQUENCE OF 36-562.  
RC TISSUE=Melanoma;  
RX MEDLINE=85000468; PubMed=6433976;  
RA Pohl G., Kællstroem M., Bergsdorf N., Wallen P., Joernvall H.;  
RT "Tissue plasminogen activator: peptide analyses confirm an indirectly  
RT derived amino acid sequence, identify the active site serine residue,  
RT establish glycosylation sites, and localize variant differences.";  
RL Biochemistry 23:3701-3707(1984).  
RN [13]  
RP SEQUENCE OF 33-52 AND 311-330.  
RC TISSUE=Melanoma;  
RX MEDLINE=83209620; PubMed=6682760;  
RA Wallen P., Pohl G., Bergsdorf N., Raanby M., Ny T., Joernvall H.;  
RT "Purification and characterization of a melanoma cell plasminogen  
RT activator.";  
RL Eur. J. Biochem. 132:681-686(1983).  
RN [14]  
RP STRUCTURE OF CARBOHYDRATES.  
RX MEDLINE=90092112; PubMed=2513186;  
RA Pfeiffer G., Schmidt M., Strube K.-H., Geyer R.;  
RT "Carbohydrate structure of recombinant human uterine tissue  
RT plasminogen activator expressed in mouse epithelial cells.";  
RL Eur. J. Biochem. 186:273-286(1989).  
RN [15]  
RP CARBOHYDRATE-LINKAGE SITE THR-96.  
RX MEDLINE=91159408; PubMed=1900431;  
RA Harris R.J., Leonard C.K., Guzzetta A.W., Spellman M.W.;  
RT "Tissue plasminogen activator has an O-linked fucose attached to  
RT threonine-61 in the epidermal growth factor domain.";  
RL Biochemistry 30:2311-2314(1991).  
RN [16]  
RP DISULFIDE BONDS IN KRINGLE 2.  
RX MEDLINE=91244765; PubMed=1645336;  
RA Vlahos C.J., Wilhelm O.G., Hassell T., Jaskunas S.R., Bang N.U.;  
RT "Disulfide pairing of the recombinant kringle-2 domain of tissue  
RT plasminogen activator produced in *Escherichia coli*.";  
RL J. Biol. Chem. 266:10070-10072(1991).  
RN [17]  
RP X-RAY CRYSTALLOGRAPHY (2.3 ÅNGSTRÖMS) OF CATALYTIC DOMAIN.  
RX MEDLINE=96200985; PubMed=8613982;  
RA Lamba D., Bauer M., Huber R., Fischer S., Rudolph R., Kohnert U.,  
RA Bode W.;  
RT "The 2.3 Å crystal structure of the catalytic domain of recombinant  
RT two-chain human tissue-type plasminogen activator.";  
RL J. Mol. Biol. 258:117-135(1996).



[18]  
RN X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF CATALYTIC DOMAIN.  
RX MEDLINE=97449126; PubMed=9305622;  
RA Renatus M., Engh R.A., Stubbs M.T., Huber R., Fischer S., Kohnert U.,  
Bode W.;  
RT "Lysine 156 promotes the anomalous proenzyme activity of tPA: X-ray  
RT crystal structure of single-chain human tPA";  
RL EMO J. 16:4797-4805(1997).  
RN [19]  
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF KRINGLE 2.  
RX MEDLINE=92118803; PubMed=1310033;  
RA de Vos A., Uitsch M.H., Kelley R.F., Padmanabhan K., Tulinsky A.,  
RA Westbrook M.L., Kossiakof A.A.;  
RT "Crystal structure of the kringle 2 domain of tissue plasminogen  
RT activator at 2.4-A resolution";  
RL Biochemistry 31:270-279(1992).  
RN [20]  
RP STRUCTURE BY NMR OF KRINGLE 2.  
RX MEDLINE=90122799; PubMed=2558718;  
RA Byeon I.-J.L., Kelley R.F., Llinas M.;  
RT "1H NMR structural characterization of a recombinant kringle 2 domain  
RT from human tissue-type plasminogen activator";  
RL Biochemistry 28:9350-9360(1989).  
RN [21]  
RP STRUCTURE BY NMR OF KRINGLE 2.  
RX MEDLINE=91200042; PubMed=1901789;  
RA Byeon I.-J.L., Kelley R.F., Llinas M.;  
RT "Kringle-2 domain of the tissue-type plasminogen activator. 1H-NMR  
RT assignments and secondary structure";  
RL Eur. J. Biochem. 197:155-165(1991).  
RN [22]  
RP STRUCTURE BY NMR OF KRINGLE 2.  
RX MEDLINE=92106329; PubMed=1762144;  
RA Byeon I.-J.L., Llinas M.;  
RT "Solution structure of the tissue-type plasminogen activator kringle  
RT 2 domain complexed to 6-aminohexanoic acid an antifibrinolytic  
RT drug";  
RL J. Mol. Biol. 222:1035-1051(1991).  
RN [23]  
Query Match 44.3%; Score 226; DB 1; Length 562;  
Best Local Similarity 47.7%; Pred. No. 1.8e-18;  
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;  
QY 2 TCYEGNHFGYKASTDTMGPCPLFWSATVLQOTYHAHRSDALQLGLKHNYCRNPDNR 61  
Db 126 TCYEDQGISYRGTVTAESGAECTNWNSSALQAQPYSGRRPDATRLGLGNHNYCRNPD 185  
QY 62 RRPWCYQVGLKPLVQECMWHDCADG 87  
Db 186 SKPWCYFKAGKYSSEFCSTACSEG 211  
RESULT 10  
TPA\_RAT  
ID TPA\_RAT STANDARD; PRT; 559 AA.  
AC P19637;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)  
DE (t-PA) (t-plasminogen activator).  
GN PLAT.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
SEQUENCE FROM N.A.  
RX MEDLINE=89170114; PubMed=3148445;  
RA Ny T., Leonardsson G., Hsueh A.J.W.;  
RT "Cloning and characterization of a cDNA for rat tissue-type  
RT plasminogen activator";

DNA 7:671-677(1988).  
[2]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=90130448; PubMed=2105315;  
RA Feng P., Ohlsson M., Ny T.;  
RT "The structure of the TATA-less rat tissue-type plasminogen activator  
RT gene. Species-specific sequence divergences in the promoter predict  
RT differences in regulation of gene expression.";  
RL J. Biol. Chem. 265:2022-2027(1990).  
CC -1- FUNCTION: Converts the abundant, but inactive, zymogen plasminogen  
CC to plasmin by hydrolyzing a single Arg-Val bond in plasminogen. By  
CC controlling plasmin-mediated proteolysis, it plays an important  
CC role in tissue remodeling and degradation, in cell migration and  
CC many other physiopathological events.  
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
CC plasminogen to form plasmin.  
CC -1- SUBUNIT: Heterodimer of chain A and chain B held by a disulfide  
CC bond.  
CC -1- SUBCELLULAR LOCATION: Secreted; extracellular.  
CC -1- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER  
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER  
CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.  
CC -1- MISCELLANEOUS: Binds to the kringle structure of the fibrin A  
CC chain. Binding to fibrin enhances its catalytic activity.  
CC -1- SIMILARITY: Belongs to peptidase family S1.  
CC -1- SIMILARITY: Contains 1 EGF-like domain.  
CC -1- SIMILARITY: Contains 1 fibronectin type I domain.  
CC -1- SIMILARITY: Contains 2 kringle domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collabora-  
CC tion between the Swiss Institute of Bioinformatics and the EMBL Outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercia  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
EMBL; M23697; AAA41812.1; -  
EMBL; M31197; AAA42261.1; -  
EMBL; M31185; AAA42261.1; JOINED.  
EMBL; M31186; AAA42261.1; JOINED.  
EMBL; M31187; AAA42261.1; JOINED.  
EMBL; M31188; AAA42261.1; JOINED.  
EMBL; M31189; AAA42261.1; JOINED.  
EMBL; M31190; AAA42261.1; JOINED.  
EMBL; M31191; AAA42261.1; JOINED.  
EMBL; M31192; AAA42261.1; JOINED.  
EMBL; M31193; AAA42261.1; JOINED.  
EMBL; M31194; AAA42261.1; JOINED.  
EMBL; M31195; AAA42261.1; JOINED.  
EMBL; M31196; AAA42261.1; JOINED.  
EMBL; A19618; CAA01482.1; -  
PIR; A35029; A35029.  
HSP; P00750; 1RTF.  
MEROPS; S01.232; -  
InterPro; IPR009003; Cys Ser trypsin.  
InterPro; IPR006209; EGF-like  
InterPro; IPR000083; Fibnctn1.  
InterPro; IPR006210; IEGF.  
InterPro; IPR000001; Kringle.  
InterPro; IPR001254; Peptidase S1.  
InterPro; IPR001314; Peptidase\_S1A.  
Pfam; PF00008; EGF; 1.  
Pfam; PF00039; fn1; 1.  
Pfam; PF00051; kringle; 2.  
Pfam; PF00089; trypsin; 1.  
PRINTS; PR00722; CHYMOTRYPSIN.  
PRINTS; PR00018; KRINGLE.  
ProDom; PD000395; Kringle; 2.  
SMART; SM00181; EGF; 1.  
SMART; SM00058; FN1; 1.  
SMART; SM00130; KR; 2.  
SMART; SM00020; Tryp\_SPC; 1.

```
DR PROSITE; PS00022; EGF_1; 1.  
DR PROSITE; EGF_2; 1.  
DR PROSITE; EGF_3; 1.  
DR PROSITE; PS00026; FIBRONECTIN_1; 1.  
DR PROSITE; PS00021; KRINGLE_1; 2.  
DR PROSITE; PS00070; TRYPSIN_DOM; 1.  
DR PROSITE; PS00240; TRYPSIN_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN_SER; 1.  
KW Plasminogen activator; Hydrolyase, Serine protease; Glycoprotein;  
KW Plasma; Kringle; EGF-like domain; Repeat; Signal.  
FT SIGNAL 1 17  
FT PROPEP 18 29  
FT CHAIN 30 559  
FT CHAIN 30 308  
FT CHAIN 309 559  
FT DOMAIN 36 78  
FT DOMAIN 79 117  
FT DOMAIN 124 205  
FT DOMAIN 213 294  
FT DOMAIN 309 559  
FT ACT_SITE 355 355  
FT ACT_SITE 404 404  
FT ACT_SITE 510 510  
FT DISULFID 38 68  
FT DISULFID 66 75  
FT DISULFID 83 94  
FT DISULFID 88 105  
FT DISULFID 107 116  
FT DISULFID 124 205  
FT DISULFID 145 187  
FT DISULFID 176 200  
FT DISULFID 213 294  
FT DISULFID 234 276  
FT DISULFID 265 289  
FT DISULFID 297 428  
FT DISULFID 340 356  
FT DISULFID 348 417  
FT DISULFID 442 516  
FT DISULFID 474 490  
FT DISULFID 506 534  
FT CARBOHYD 149 149  
FT CARBOHYD 481 481  
FT CONFLICT 380 380  
SQ SEQUENCE 559 AA; 62903 MW; 7D5D3809C1D1C921 CRC64;  
  
Query Match 44.1%; Score 225; DB 1; Length 559;  
Best Local Similarity 47.7%; Pred. No. 2.4e-18;  
Matches 41; Conservative 10; Mismatches 35; Indels 0; Gaps 0;  
  
Qy 2 TCYEGNGHYFGKASTDTMGPCLPWNSATVLYQTYHAHSDALQLGLGKHNYCRNPDR 61  
Db 123 TCYEGNGHYFGKASTDTMGPCLPWNSATVLYQTYHAHSDALQLGLGKHNYCRNPDR 182  
Qy 62 RRPACVYQVGLKPLVQCMWHDGADG 87  
Db 183 VKPCVYFKACKYTFECSTPACPKG 208  
  
RESULT 11  
URTG DESRO STANDARD; PRT; 394 AA.  
AC P49150;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Salivary plasminogen activator gamma precursor (EC 3.4.21.68) (DSPA  
DE gamma).  
OS Desmodus rotundus (Vampire bat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
```

FT DISULFID 308 324 BY SIMILARITY.  
FT DISULFID 341 369 BY SIMILARITY.  
FT CARBOHYD 315 N-LINKED (GLCNAc...) (POTENTIAL).  
SQ SEQUENCE 394 AA; 44105 MW; 90CD6F52F3D19FDC CRC64;  
  
Query Match 43.1%; Score 220; DB 1; Length 394;  
Best Local Similarity 45.2%; Pred. No. 6.2e-18;  
Matches 38; Conservative 13; Mismatches 33; Indels 0; Gaps 0;  
  
QY 2 TCYEGNHFGKASTDMGRCPTWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDR 61  
DQ 44 TCYKDGQVYRGTWTSTSGAGCINWNSNLLIRTYNGRMPKAVLGLGHNHNYCRNPDA 103  
QY 62 RRPWCYVQVGLKPLVQECWHDCA 85  
DQ 104 SKPWCYVIKARFTSSECSVPVCS 127  
  
RESULT 12  
TPA\_MOUSE  
ID TPA\_MOUSE STANDARD; PRT; 559 AA.  
AC P11214; Q91VP2;  
DT 01-JUL-1989 (Rel. 11, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tpa)  
DE (t-pa) (t-plasminogen activator).  
GN PLAT.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_taxonomy:10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88087303; PubMed=2826484;  
RA Rickles R.J., Darrow A.L., Strickland S.;  
RT "Molecular cloning of complementary DNA to mouse tissue plasminogen  
RT activator mRNA and its expression during F9 teratocarcinoma cell  
RT differentiation.";  
RL J. Biol. Chem. 263:1563-1569(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Mammary gland;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Touchman J.W., Green E.D., Dickinson M.C.,  
RA Rodriquez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.T., Skalska U., Smalus D.B.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -!- FUNCTION: Converts the abundant, but inactive, zymogen plasminogen  
CC to plasmin by hydrolyzing a single Arg-Val bond in plasminogen. By  
CC controlling plasmin-mediated proteolysis, it plays an important  
CC role in tissue remodeling and degradation, in cell migration and  
CC many other physiological events.  
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
CC plasminogen to form plasmin.  
CC -!- SUBUNIT: Heterodimer of chain A and chain B held by a disulfide  
CC bond.

CC -!- SUBCELLULAR LOCATION: Secreted; extracellular.  
CC -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER  
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER  
CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.  
CC -!- MISCELLANEOUS: Binds to the kringle structure of the fibrin A  
CC chain. Binding to fibrin enhances its catalytic activity.  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
CC -!- SIMILARITY: Contains 1 EGF-like domain. I domain.  
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.  
CC -!- SIMILARITY: Contains 2 kringle domains.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboratio  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on it  
CC use by non-profit institutions as long as its content is in no wa  
CC modified and this statement is not removed. Usage by and for commercia  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL; J03520; AAA40470.1; -;  
CC EMBL; BC011256; AAH11256.1; -;  
CC PIR; A29941; A29941.  
CC KSSP; P00750; IASH.  
CC MEROPS; S01.232; -;  
CC MGD; MGI:97610; Plat.  
CC InterPro; IPR009003; Cys Ser trypsin.  
CC InterPro; IPR006209; EGF-like.  
CC InterPro; IPR000083; Fibrinectn.  
CC InterPro; IPR006210; IEGF.  
CC InterPro; IPR000001; Kringle.  
CC InterPro; IPR001254; Peptidase\_S1.  
CC InterPro; IPR001314; Peptidase\_S1A.  
CC Pfam; PF00008; EGF; 1.  
CC Pfam; PF00039; fn1; 1.  
CC Pfam; PF00051; kringle; 2.  
CC Pfam; PF00089; trypsin; 1.  
CC PRINTS; PR00722; CHYMOTRYPSIN.  
CC PRINTS; PR00018; KRINGLE.  
CC ProDom; PD000395; Kringle; 2.  
CC SMART; SM00181; EGF; 1.  
CC SMART; SM00058; FN1; 1.  
CC SMART; SM00130; KR; 2.  
CC SMART; SM00020; TRYD\_SPC; 1.  
CC PROSITE; PS00022; EGF\_1; 1.  
CC PROSITE; PS01186; EGF\_2; 1.  
CC PROSITE; PS50026; EGF\_3; 1.  
CC PROSITE; PS01253; FIBRONECTIN\_1; 1.  
CC PROSITE; PS00021; KRINGLE\_1; 2.  
CC PROSITE; PS50070; KRINGLE\_2; 2.  
CC PROSITE; PS50240; TRYPSIN\_DOM; 1.  
CC PROSITE; PS00134; TRYPSIN\_HIS; 1.  
CC PROSITE; PS00135; TRYPSIN\_SER; 1.  
CC Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
CC Plasma; Kringle; EGF-like domain; Repeat; Signal.  
CC SIGNAL 1 17 PROBABLE.  
FT PROPEP 18 29 TISSUE-TYPE PLASMINOGEN ACTIVATOR.  
FT CHAIN 30 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR A  
FT CHAIN 30 308 CHAIN.  
FT CHAIN 309 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR B  
FT CHAIN 309 559 CHAIN.  
FT DOMAIN 36 78 FIBRONECTIN TYPE-I.  
FT DOMAIN 79 117 EGF-LIKE.  
FT DOMAIN 124 205 KRINGLE 1.  
FT DOMAIN 213 294 KRINGLE 2.  
FT DOMAIN 309 559 SERINE PROTEASE  
FT ACT\_SITE 355 355 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 404 404 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 510 510 CHARGE RELAY SYSTEM.  
FT DISULFID 38 68 BY SIMILARITY.  
FT DISULFID 66 75 BY SIMILARITY.  
FT DISULFID 83 94 BY SIMILARITY.  
FT DISULFID 88 105 BY SIMILARITY.

```
FT DISULFID 107 116 BY SIMILARITY.
FT DISULFID 124 205 BY SIMILARITY.
FT DISULFID 145 187 BY SIMILARITY.
FT DISULFID 176 200 BY SIMILARITY.
FT DISULFID 213 234 BY SIMILARITY.
FT DISULFID 234 276 BY SIMILARITY.
FT DISULFID 265 289 BY SIMILARITY.
FT DISULFID 297 428 INTERCHAIN (BY SIMILARITY).
FT DISULFID 340 356 BY SIMILARITY.
FT DISULFID 348 417 BY SIMILARITY.
FT DISULFID 442 516 BY SIMILARITY.
FT DISULFID 474 490 BY SIMILARITY.
FT DISULFID 506 534 BY SIMILARITY.
FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 260 260 G -> A (IN REF. 1).
FT CONFLICT 325 325 P -> A (IN REF. 1).
SQ SEQUENCE 559 AA; 63122 MW; 8CCE2BDB94514D9 CRC64;

Query Match 42.9%; Score 219; DB 1; Length 559;
Best Local Similarity 46.0%; Pred. No. 1.2e-17;
Matches 40; Conservative 11; Mismatches 36; Indels 0; Gaps 0;

OY 2 TCYEGNGHFYKSTDMGRCPLPNSATVLTQTYHAHRSDALQLGLGHNYCNEPDR 61
Db 123 TCFEQGYTYGTWSTABSGACINWNSVLSLPYNARRNPAIKLGLGHNYCNEPDR 182

OY 62 RRPWCYQVGLKPLVQECMHDCADGK 88
Db 183 LKPWCYVFKAGKYTEFCSTACPCKGK 209

RESULT 13
URTL_DESRO STANDARD; PRT; 477 AA.
AC P98119;
AD 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE Salivary plasminogen activator alpha 1 precursor (EC 3.4.21.66) (DSPA
DE alpha-1).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: Cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RC TISSUE=Salivary gland;
RX MEDLINE=98022741; PubMed=9354616;
RA Renatus M., Stubbs M.T., Huber R., Bringmann P., Donner P.,
RA Schleuning W.D., Bode W.;
RT "Catalytic domain structure of vampire bat plasminogen activator: a
RT molecular paradigm for proteolysis without activation cleavage.";
RL Biochemistry 36:13483-13493(1997).
```

```
CC -I- FUNCTION: Probably essential to support the feeding habits of this
CC exclusively haematophagous animal. Potent thrombolytic agent.
CC -I- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -I- ENZYME REGULATION: Activity toward plasminogen is stimulated in
CC the presence of fibrin I.
CC -I- SUBUNIT: Monomer.
CC -I- DOMAIN: The fibronectin type-I domain mediates binding to fibrin,
CC and the kringle domain apparently mediates fibrin-induced
CC stimulation of activity.
CC -I- SIMILARITY: Belongs to peptidase family S1.
CC -I- SIMILARITY: Contains 1 EGF-like domain.
CC -I- SIMILARITY: Contains 1 fibronectin type I domain.
CC -I- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collabora-
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercia
CC entities requires a license agreement (See http://www.isb-sib.ch/announce
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M63987; AAA1591.1; -
CC EMBL; M63986; AAA1592.1; -
CC PIR; JS0597; JS0597.
CC PDB; 1A51; 23-MAR-99.
CC MEROPS; S01.232; -
CC GlycoSuiteDB; P98119; -
CC InterPro; IPR009003; Cys Ser trypsin.
CC InterPro; IPR006209; EGF-like.
CC InterPro; IPR000083; Fibrinctnl.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001314; Peptidase_S1A.
CC InterPro; IPR001254; Peptidase_S1.
CC Pfam; PF00008; EGF_1.
CC Pfam; PF00039; fni; 1.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00181; EGF; 1.
CC SMART; SM00058; FNI; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00020; Tryp_Spc; 1.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS50026; EGF_3; 1.
CC PROSITE; PS01253; FIBRONECTIN_1; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS50070; KRINGLE_2; 1.
CC PROSITE; PS50240; TRYPSIN_DOM; 1.
CC PROSITE; PS50134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
CC Kringle; EGF-like domain; Signal; Multigene family; 3D-structure.
FT SIGNAL 1 36 POTENTIAL PLASMINOGEN ACTIVATOR ALPHA 1.
FT CHAIN 37 477 SALIVARY PLASMINOGEN ACTIVATOR ALPHA 1.
FT DOMAIN 40 82 FIBRONECTIN TYPE-I.
FT DOMAIN 83 121 EGF-LIKE.
FT DOMAIN 128 209 KRINGLE.
FT DOMAIN 225 477 SERINE PROTEASE.
FT ACT_SITE 272 272 CHARGE RELAY SYSTEM.
FT ACT_SITE 321 321 CHARGE RELAY SYSTEM.
FT ACT_SITE 428 428 CHARGE RELAY SYSTEM.
FT DISULFID 42 72 BY SIMILARITY.
FT DISULFID 70 79 BY SIMILARITY.
FT DISULFID 87 98 BY SIMILARITY.
FT DISULFID 92 109 BY SIMILARITY.
FT DISULFID 111 120 BY SIMILARITY.
FT DISULFID 128 209 BY SIMILARITY.
```



DR PROSITE; PS50240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Plasmagen activation; Hydrolyase; Serine protease; Glycoprotein;  
FT SIGNAL; Kringle; EGF-like domain; Repeat; Signal.  
FT PROPEL 1 21  
FT CHAIN 22 33  
FT CHAIN 34 566  
FT CHAIN 34 314  
FT CHAIN 315 566  
FT DOMAIN 40 82  
FT DOMAIN 83 121  
FT DOMAIN 128 209  
FT DOMAIN 219 300  
FT DOMAIN 315 566  
FT ACT\_SITE 361 361  
FT ACT\_SITE 410 410  
FT ACT\_SITE 517 517  
FT DISULFID 42 72  
FT DISULFID 70 79  
FT DISULFID 87 98  
FT DISULFID 92 109  
FT DISULFID 111 120  
FT DISULFID 128 209  
FT DISULFID 149 191  
FT DISULFID 180 204  
FT DISULFID 219 300  
FT DISULFID 240 282  
FT DISULFID 271 295  
FT DISULFID 303 434  
FT DISULFID 346 362  
FT DISULFID 354 423  
FT DISULFID 448 523  
FT DISULFID 480 496  
FT DISULFID 513 541  
FT CARBOHYD 153 153  
FT CARBOHYD 487 487  
SQ SEQUENCE 566 AA; 63701 MW; 2EB6BBE4E32276C3 CRC64;

Query Match 41.0%; Score 209; DB 1; Length 566;  
Best Local Similarity 47.6%; Pred No. 1.7e-16;  
Matches 39; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

OY 3 CYEGNGHYRGKASTDTWGRPLPWSATVLTQTYHARSDALQGLGKHYCRPNRR 62  
DB 219 CYTGNGLAYRGTRSHKSGASCLPWSVFLTSKIYTAWSKAPALGLGKHCRNPDDA 278  
OY 63 RPWCYVQVGLKPLVQECMVHDC 84  
DB 279 QPWCHVWXRQLTWBYCDVPOC 300

RESULT 15  
UROK\_CHICK STANDARD; PRT; 434 AA.  
AC P15120;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)  
DE (U-plasminogen activator).  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90110185; PubMed=2295632;  
RA Leslie N.D., Kessler C.A., Bell S.M., Degen J.L.;  
RI "The chicken urokinase-type plasminogen activator gene."

J. Biol. Chem. 265:1339-1344(1990).  
-1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
plasmagen to form plasmin.  
-1- SIMILARITY: Belongs to peptidase family S1.  
-1- SIMILARITY: Contains 1 EGF-like domain.  
-1- SIMILARITY: Contains 1 kringle domain.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboratio  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
at the European Bioinformatics Institute. There are no restrictions on it  
use by non-profit institutions as long as its content is in no wa  
modified and this statement is not removed. Usage by and for commercia  
entities requires a license agreement (See http://www.isb-sib.ch/announce  
or send an email to license@isb-sib.ch).  
-----  
EMBL; J05187; AAA49131.1; -  
EMBL; J05188; AAA49130.1; -  
PIR; A35005; A35005.  
HSP; P00763; IDPO.  
MROPS; S01.231; -  
InterPro; IPR009003; Cys Ser trypsin.  
InterPro; IPR006209; EGF\_like.  
InterPro; IPR006210; IEGF.  
InterPro; IPR000001; Kringle.  
InterPro; IPR008293; Peptidase S1.  
InterPro; IPR001254; Peptidase S1.  
InterPro; IPR001314; Peptidase\_S1A.  
Pfam; PF00051; kringle; 1.  
Pfam; PF00089; trypsin; 1.  
PIRSE; PIRSF001144; UTK Plas act; 1.  
PRINTS; PR00722; CHYMOTRYPSIN.  
PRINTS; PR00018; KRINGLE.  
ProDom; PD000395; Kringle; 1.  
SMART; SM00181; EGF; 1.  
SMART; SM00130; KR; 1.  
SMART; SM00020; Tryp\_Spc; 1.  
PROSITE; PS00022; EGF\_1; 1.  
PROSITE; PS01186; EGF\_2; 1.  
PROSITE; PS00026; EGF\_3; 1.  
PROSITE; PS00021; KRINGLE 1; 1.  
PROSITE; PS00070; KRINGLE 2; 1.  
PROSITE; PS00240; TRYPSIN\_DOM; 1.  
PROSITE; PS00134; TRYPSIN\_HIS; 1.  
PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Plasmagen activation; Hydrolyase; Serine protease; Glycoprotein;  
KW Kringle; EGF-like domain; Signal; zymogen.  
FT SIGNAL 1 230 POTENTIAL.  
FT CHAIN 21 434 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.  
FT CHAIN 21 171 CHAIN A (BY SIMILARITY).  
FT CHAIN 173 434 CHAIN B (BY SIMILARITY).  
FT DOMAIN 36 72 EGF-LIKE.  
FT DOMAIN 79 158 KRINGLE.  
FT DOMAIN 159 172 CONNECTING PEPTIDE.  
FT DOMAIN 173 434 SERINE PROTEASE.  
FT DISULFID 40 48 BY SIMILARITY.  
FT DISULFID 42 60 BY SIMILARITY.  
FT DISULFID 62 71 BY SIMILARITY.  
FT DISULFID 162 296 INTERCHAIN (BY SIMILARITY).  
FT DISULFID 202 218 BY SIMILARITY.  
FT DISULFID 210 285 BY SIMILARITY.  
FT DISULFID 310 379 BY SIMILARITY.  
FT DISULFID 342 358 BY SIMILARITY.  
FT DISULFID 369 397 BY SIMILARITY.  
FT ACT\_SITE 217 217 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 272 272 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 373 373 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT CARBOHYD 228 228 N-LINKED (GLCNAC...) (POTENTIAL).  
SQ SEQUENCE 434 AA; 49400 MW; BD801048DD666A55 CRC64;  
Query Match 38.9%; Score 198.5; DB 1; Length 434;  
Best Local Similarity 54.4%; Pred. No. 2.1e-15;  
Matches 37; Conservative 7; Mismatches 19; Indels 5; Gaps 2

QY 3 CYEGNGHYRGAHSTDTWGRPCLPWNSATVLO-QYVHAHRSALQLGLGKKNYCRNPDNR 61  
DB 79 CYSNGEDYRGMAEDP-GLYWDHPSVIRMGDYLADLKNALQLGLGKKNYCRNPNR 134  
QY 62 RRPWCYVO 69  
DB 135 SRPWCYTK 142

RESULT 16  
HGFA HUMAN STANDARD; PRT; 655 AA.  
AC Q04756; Q14726;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hepatocyte growth factor activator precursor (BC 3.4.21.-) (HGF  
activator) (HGFA).  
GN HGFA.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Liver, and Serum;  
RX MEDLINE=93252878; PubMed=7683665;  
RA Miyazawa K., Shimomura T., Kitamura A., Kondo J., Morimoto Y.,  
RA Kitamura N.;  
RT "Molecular cloning and sequence analysis of the cDNA for a human  
RT serine protease responsible for activation of hepatocyte growth  
RT factor. Structural similarity of the protease precursor to blood  
RT coagulation factor XII.";  
RL J. Biol. Chem. 268:10024-10028 (1993).  
RN [2]  
RP SEQUENCE OF 40-655 FROM N.A.  
RA Zhao S., Odell C.;  
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Activates hepatocyte growth factor (HGF) by  
CC converting it from a single chain to a heterodimeric form.  
CC -!- SUBUNIT: Dimer of a short chain and a long chain linked by a  
CC disulfide bond.  
CC -!- SUBCELLULAR LOCATION: Secreted as an inactive single-chain  
CC precursor and is then activated to a heterodimeric form.  
CC -!- TISSUE SPECIFICITY: Liver.  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
CC -!- SIMILARITY: Contains 2 EGF-like domains.  
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.  
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.  
CC -!- SIMILARITY: Contains 1 kringle domain.  
CC -!- CAUTION: It is uncertain whether Met-1 is the initiator.  
CC  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; D14012; BAA03113.1;  
CC EMBL; Z69923; -; NOT\_ANNOTATED\_CDS.  
CC PIR; A46688; A46688.  
CC HSP; P00763; LDPO.  
CC MEROPS; S01.228; -;  
CC Genew; HGNC:4894; HGFA.  
CC MIM; 604552; -;  
CC GO; GO:0005576; C:extracellular; TAS.  
CC GO; GO:004252; F:serine-type endopeptidase activity; TAS.  
CC GO; GO:0046508; P:proteolysis and peptidolysis; TAS.  
CC InterPro; IPR009003; Cys\_Ser\_trypsin.  
CC InterPro; IPR000742; EGF\_2.  
CC InterPro; IPR006209; EGF\_like.

InterPro; IPR000083; Fibrinctnl.  
InterPro; IPR000562; FN\_Type\_II.  
InterPro; IPR006210; IEGF.  
InterPro; IPR000001; Kringle.  
InterPro; IPR001254; Peptidase\_S1.  
InterPro; IPR001314; Peptidase\_S1A.  
Pfam; PF00008; EGF\_2.  
Pfam; PF00039; fn1; 1.  
Pfam; PF00040; fn2; 1.  
Pfam; PF00051; kringle; 1.  
Pfam; PF00089; trypsin; 1.  
PRINTS; PR00722; CHYMOTRYPSIN.  
PRINTS; PR00013; ENTPEPIL.  
PRINTS; PR00018; KRINGLE.  
ProDom; PD000995; FN\_Type\_II; 1.  
ProDom; PD000395; Kringle; 1.  
SMART; SM00181; EGF; 2.  
SMART; SM00058; FN1; 1.  
SMART; SM00059; FN2; 1.  
SMART; SM00130; KR; 1.  
SMART; SM00020; Tryp\_Spc; 1.  
PROSITE; PS00022; EGF\_1; 2.  
PROSITE; PS01186; EGF\_2; 1.  
PROSITE; PS50026; EGF\_3; 2.  
PROSITE; PS01253; FIBRONECTIN\_1; 1.  
PROSITE; PS00023; FIBRONECTIN\_2; 1.  
PROSITE; PS00021; KRINGLE\_1; 1.  
PROSITE; PS00070; KRINGLE\_2; 1.  
PROSITE; PS50240; TRYPSIN\_DOM; 1.  
PROSITE; PS00134; TRYPSIN\_HIS; 1.  
PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;  
EGF-like domain; Repeat; Zymogen.  
FT SIGNAL 1 30  
FT PROPEP 31 372  
FT CHAIN 373 407  
FT  
FT CHAIN 408 655  
FT  
FT DOMAIN 108 148  
FT DOMAIN 160 198  
FT DOMAIN 200 240  
FT DOMAIN 241 279  
FT DOMAIN 286 367  
FT DOMAIN 408 655  
FT ACT\_SITE 447 447  
FT ACT\_SITE 497 497  
FT ACT\_SITE 598 598  
FT DISULFID 108 133  
FT DISULFID 122 148  
FT DISULFID 164 175  
FT DISULFID 169 186  
FT DISULFID 188 197  
FT DISULFID 202 230  
FT DISULFID 228 237  
FT DISULFID 245 256  
FT DISULFID 250 267  
FT DISULFID 269 278  
FT DISULFID 286 367  
FT DISULFID 307 349  
FT DISULFID 338 362  
FT DISULFID 394 521  
FT DISULFID 432 448  
FT DISULFID 440 510  
FT DISULFID 535 604  
FT DISULFID 567 583  
FT DISULFID 594 622  
FT CARBOHYD 48 48  
FT CARBOHYD 290 290  
FT CARBOHYD 468 468  
FT CARBOHYD 492 492  
FT CARBOHYD 546 546  
FT CONFLICT 644 644

CLEAVED IN ACTIVE FORM.  
HEPATOCYTE GROWTH FACTOR ACTIVATOR: SHORT  
CHAIN.  
HEPATOCYTE GROWTH FACTOR ACTIVATOR: LONG  
CHAIN.  
FIBRONECTIN TYPE-II.  
EGF-LIKE 1.  
FIBRONECTIN TYPE-I.  
EGF-LIKE 2.  
KRINGLE.  
SERINE PROTEASE.  
CHARGE RELAY SYSTEM (BY SIMILARITY).  
CHARGE RELAY SYSTEM (BY SIMILARITY).  
CHARGE RELAY SYSTEM (BY SIMILARITY).  
BY SIMILARITY.  
BY SIMILARITY.  
BY SIMILARITY.  
BY SIMILARITY.  
BY SIMILARITY.  
BY SIMILARITY.  
BY SIMILARITY.  
INTERCHAIN (BY SIMILARITY).  
BY SIMILARITY.  
BY SIMILARITY.  
BY SIMILARITY.  
BY SIMILARITY.  
BY SIMILARITY.  
N-LINKED (GLCNAC. . .)  
N-LINKED (GLCNAC. . .)  
N-LINKED (GLCNAC. . .)  
N-LINKED (GLCNAC. . .)  
N-LINKED (GLCNAC. . .)  
R -> Q (IN REF. 2).



SQ SEQUENCE 655 AA; 70681 MW; 2CF72F1EB862ED7 CRC64;

Query Match 37.8%; Score 193; DB 1; Length 655;  
 Best Local Similarity 54.5%; Pred. No. 1.4e-14;  
 Matches 36; Conservative 4; Mismatches 26; Indels 0; Gaps 0;

QY 3 CYEGNGHYRGKASDTWGRPCLPNHSATVLCQTVHAHRSALGLGKKNHCNPNRR 62  
 286 CFLNGTGYRGVASTASGLSLCLAWNSDLLYQELHVDVSGAALLGLGPHAYCRPNDE 345  
 DQ  
 QY 63 RPWCYV 68  
 DQ  
 DB 346 RPWCYV 351

RESULT 17

HGFA MOUSE STANDARD; PRT; 653 AA.

AC Q9R058; Q9JKV4;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF  
 DE activator) (HGFA).  
 GN HGFA.

OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10990;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c;  
 RA Itoh H., Kataoka H., Koono H.;  
 RT "Mouse hepatocyte growth factor activator."  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21226753; PubMed=11032833;  
 RA van Adelsberg J.S., Sehgal S., Kukes A., Brady C., Barasch J.,  
 RA Yang J., Huan Y.;  
 RT "Activation of hepatocyte growth factor (HGF) by endogenous HGF  
 RT activator is required for metanephric kidney morphogenesis in  
 RT vitro".  
 RL J. Biol. Chem. 276:15099-15106(2001).  
 CC -!- FUNCTION: Activates hepatocyte growth factor (HGF) by converting  
 CC it from a single chain to a heterodimeric form (By similarity).  
 CC -!- SUBUNIT: Dimer of a short chain and a long chain linked by a  
 CC disulfide bond (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Secreted as an inactive single-chain  
 CC precursor and is then activated to a heterodimeric form (By  
 CC similarity).  
 CC -!- SIMILARITY: Belongs to peptidase family S1.  
 CC -!- SIMILARITY: Contains 2 EGF-like domains.  
 CC -!- SIMILARITY: Contains 1 fibronectin type I domain.  
 CC -!- SIMILARITY: Contains 1 fibronectin type II domain.  
 CC -!- SIMILARITY: Contains 1 kringle domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF099017; AAF02489.1; -;  
 CC DR EMBL; AF224724; AAF34712.1; -;  
 CC DR HSSP; P00763; IDPO.  
 CC DR MEROPS; S01.228; -;  
 CC DR MGD; MG1:1859281; Hgfac.  
 CC DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
 CC DR InterPro; IPR000742; EGF 2.  
 CC DR InterPro; IPR006209; EGF like.

DR InterPro; IPR000083; Fibrinctnl.  
 DR InterPro; IPR000562; FN\_Type\_II.  
 DR InterPro; IPR006210; IEFG.  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR Pfam; PF00008; EGF\_2.  
 DR Pfam; PF00039; fn1; 1.  
 DR Pfam; PF00040; fn2; 1.  
 DR Pfam; PF00051; kringle; 1.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PRINTS; PR00013; ENTPEPII.  
 DR PRINTS; PR00018; KRINGLE.  
 DR ProDom; PD000935; FN\_Type\_II; 1.  
 DR ProDom; PD000395; Kringle; 1.  
 DR SMART; SM00181; EGF\_2; 1.  
 DR SMART; SM00059; FN2; 1.  
 DR SMART; SM00130; KR; 1.  
 DR SMART; SM00020; TYP\_SPC; 1.  
 DR PROSITE; PS00022; EGF\_1; 2.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS00026; EGF\_3; 2.  
 DR PROSITE; PS01253; FIBRONECTIN\_1; 1.  
 DR PROSITE; PS00023; FIBRONECTIN\_2; 1.  
 DR PROSITE; PS00021; KRINGLE\_1; 1.  
 DR PROSITE; PS00070; KRINGLE\_2; 1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 DR Hydrolase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;  
 EGF-like domain; Repeat; Zymogen.  
 FT SIGNAL 1 29 BY SIMILARITY.  
 FT PROPEP 30 369 CLEAVED IN ACTIVE FORM (BY SIMILARITY).  
 FT CHAIN 370 405 HEPATOCYTE GROWTH FACTOR ACTIVATOR SHORT  
 FT CHAIN 406 653 HEPATOCYTE GROWTH FACTOR ACTIVATOR LONG  
 FT DOMAIN 105 145 FIBRONECTIN TYPE-II.  
 FT DOMAIN 157 195 EGF-LIKE 1.  
 FT DOMAIN 197 237 FIBRONECTIN TYPE-I.  
 FT DOMAIN 238 276 EGF-LIKE 2.  
 FT DOMAIN 283 364 KRINGLE.  
 FT DOMAIN 406 653 SERINE PROTEASE.  
 FT ACT\_SITE 445 445 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 495 495 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 596 596 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT DISULFID 105 130 BY SIMILARITY.  
 FT DISULFID 119 145 BY SIMILARITY.  
 FT DISULFID 161 172 BY SIMILARITY.  
 FT DISULFID 166 183 BY SIMILARITY.  
 FT DISULFID 185 194 BY SIMILARITY.  
 FT DISULFID 199 227 BY SIMILARITY.  
 FT DISULFID 225 234 BY SIMILARITY.  
 FT DISULFID 242 253 BY SIMILARITY.  
 FT DISULFID 247 264 BY SIMILARITY.  
 FT DISULFID 266 275 BY SIMILARITY.  
 FT DISULFID 283 364 BY SIMILARITY.  
 FT DISULFID 304 346 BY SIMILARITY.  
 FT DISULFID 335 359 BY SIMILARITY.  
 FT DISULFID 392 519 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 430 445 BY SIMILARITY.  
 FT DISULFID 438 508 BY SIMILARITY.  
 FT DISULFID 533 602 BY SIMILARITY.  
 FT DISULFID 565 581 BY SIMILARITY.  
 FT DISULFID 592 620 BY SIMILARITY.  
 FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 466 466 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 544 544 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 164 164 G -> W (IN REF. 2).

SQ SEQUENCE 653 AA; 70567 MW; 88B4B20255DF7FDC CRC64;

Query Match 36.9%; Score 188; DB 1; Length 653;  
 Best Local Similarity 53.0%; Pred. No. 5.1e-14;  
 Matches 35; Conservative 6; Mismatches 25; Indels 0; Gaps 0;

OY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTHAHRSDALQGLGKHNYCRNPDR 62  
 Db 283 CFLNGTEYRGVASTAASGLSLAWNSDLLYQELHYDSVAAAAILGLGFHAYCRNPDKR 342

OY 63 RPWCYV 68  
 Db 343 RPWCYV 348

RESULT 18  
 ID FA12\_CAVPO STANDARD; PRT; 603 AA.  
 AC Q04962;  
 DT 01-PEB-1996 (Rel. 33, Created)  
 DT 01-PEB-1996 (Rel. 33, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)  
 DE (HAF) (Fragment).  
 GN F12.  
 OS Cavia porcellus (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Hystriognathi; Cavidae; Cavia.  
 OX NCBI\_TaxId:10141;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-37; 318-332 AND 359-373.  
 RC TISSUE=Liver;  
 RX MEDLINE=93003367; PubMed=1390917;  
 RA Semba U., Yamamoto T., Kunisada T., Shibuya Y., Tanase S.,  
 RA Kambara T., Okabe H.;  
 RT "Primary structure of guinea-pig Hageman factor: sequence around the  
 RT cleavage site differs from the human molecule.";  
 RL Biochim. Biophys. Acta 1159:113-121 (1992).  
 CC -!- FUNCTION: Factor XII is a serum glycoprotein that participates in  
 CC the initiation of blood coagulation, fibrinolysis, and the  
 CC generation of bradykinin and angiotensin.  
 CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Ile bonds in factor  
 CC VII to form factor VIIa and factor XI to form factor Xla.  
 CC -!- MISCELLANEOUS: Factor XII, prekallikrein, and HMW kininogen form a  
 CC complex bound to an anionic surface. Prekallikrein is cleaved by  
 CC factor XII to form kallikrein, which then cleaves factor XII first  
 CC to alpha-factor XIIa and then to beta-factor XIIa. Alpha-factor  
 CC XIIa activates factor XI to factor Xla.  
 CC -!- SIMILARITY: Belongs to peptidase family S1.  
 CC -!- SIMILARITY: Contains 2 EGF-like domains.  
 CC -!- SIMILARITY: Contains 1 fibronectin type I domain.  
 CC -!- SIMILARITY: Contains 1 fibronectin type II domain.  
 CC -!- SIMILARITY: Contains 1 kringle domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X68615; CAA48600.1; --  
 DR PIR; S28941; S28941.  
 DR HSSP; P00763; LDPO.  
 DR MEROPS; S01.211; --  
 DR InterPro; IPR009003; Cys Ser\_trypsin.  
 DR InterPro; IPR000742; EGF 2.  
 DR InterPro; IPR001881; EGF Ca.  
 DR InterPro; IPR006209; EGF-like.  
 DR InterPro; IPR000083; Fibrnctnl.  
 DR InterPro; IPR000562; FN\_Type\_II.  
 DR InterPro; IPR006210; IEGF.

DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR001254; Peptidase S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR Pfam; PF00008; EGF 2.  
 DR Pfam; PF00039; fn1; 1.  
 DR Pfam; PF00040; fn2; 1.  
 DR Pfam; PF00051; kringle; 1.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PRINTS; PR00013; FNTYPEII.  
 DR PRINTS; PR00018; KRINGLE.  
 DR ProDom; PD000995; FN\_Type\_II; 1.  
 DR ProDom; PD000395; Kringle; 1.  
 DR SMART; SM00181; EGF 2.  
 DR SMART; SM00058; FN1; 1.  
 DR SMART; SM00059; FN2; 1.  
 DR SMART; SM00130; KR; 1.  
 DR SMART; SM00020; TRYD\_SPC; 1.  
 DR PROSITE; PS00022; EGF\_1; 2.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS00026; EGF\_3; 2.  
 DR PROSITE; PS01253; FIBRONECTIN 1; 1.  
 DR PROSITE; PS00023; FIBRONECTIN 2; 1.  
 DR PROSITE; PS00021; KRINGLE 1; 1.  
 DR PROSITE; PS00070; KRINGLE 2; 1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 DR KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;  
 KW Hydrolyase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.  
 FT NON\_TER 1 1  
 FT SIGNAL <1 18  
 FT CHAIN 19 358 ALPHA-FACTOR XIIA HEAVY CHAIN.  
 FT CHAIN 359 603 ALPHA-FACTOR XIIA LIGHT CHAIN.  
 FT DOMAIN 46 87 FIBRONECTIN TYPE-II.  
 FT DOMAIN 93 130 EGF-LIKE 1.  
 FT DOMAIN 132 172 FIBRONECTIN TYPE-I.  
 FT DOMAIN 173 209 EGF-LIKE 2.  
 FT DOMAIN 216 294 KRINGLE.  
 FT DOMAIN 312 342 PRO-RICH.  
 FT DOMAIN 359 603 SERINE PROTEASE.  
 FT ACT\_SITE 398 398 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 447 447 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 551 551 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT DISULFID 97 109 BY SIMILARITY.  
 FT DISULFID 103 118 BY SIMILARITY.  
 FT DISULFID 120 129 BY SIMILARITY.  
 FT DISULFID 134 162 BY SIMILARITY.  
 FT DISULFID 160 169 BY SIMILARITY.  
 FT DISULFID 177 188 BY SIMILARITY.  
 FT DISULFID 182 197 BY SIMILARITY.  
 FT DISULFID 199 208 BY SIMILARITY.  
 FT DISULFID 216 294 BY SIMILARITY.  
 FT DISULFID 237 276 BY SIMILARITY.  
 FT DISULFID 265 289 BY SIMILARITY.  
 FT DISULFID 345 472 BY SIMILARITY.  
 FT DISULFID 383 399 BY SIMILARITY.  
 FT DISULFID 391 461 BY SIMILARITY.  
 FT DISULFID 422 425 BY SIMILARITY.  
 FT DISULFID 488 557 BY SIMILARITY.  
 FT DISULFID 520 536 BY SIMILARITY.  
 FT DISULFID 547 578 BY SIMILARITY.  
 FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 270 270 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 603 AA; 66795 MW; 48DC6B946FB9ED59 CRC64;

Query Match 35.4%; Score 180.5; DB 1; Length 603;  
 Best Local Similarity 42.9%; Pred. No. 3.5e-13;  
 Matches 36; Conservative 11; Mismatches 32; Indels 5; Gaps 2

OY 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTHAHRSD-ALQGLGKHNYCRNPDR 60  
 Db 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTHAHRSD-ALQGLGKHNYCRNPDR 60

Db 215 SCYERGVSVYGMARTVTSAGKQWAS-...EATYRMTAEQALRRGLGHTFCRPNFN 270  
Qy 61 RRRPWCYVQVGLKPLVQCMVHDC 84  
Db 271 DTRPWCFCVWGNLSWEYCDLAQC 294

RESULT 19  
FA12 HUMAN  
ID FA12 HUMAN STANDARD; PRT; 615 AA.  
AC P00748; P78339;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)  
DE (HAP).  
GN F12.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88007593; PubMed=2888762;  
RA Cool D.E., McGillivray R.T.A.;  
RT "Characterization of the human blood coagulation factor XII gene.  
RT Intron/exon gene organization and analysis of the 5'-flanking  
RT region.";  
RL J. Biol. Chem. 262:13662-13673(1987).  
RN [2]  
RP SEQUENCE FROM N.A., AND VARIANTS ALA-207; ASP-545 AND HIS-605.  
RA Rieder M.J., Armet T.Z., Carrington D.P., Ozuna M., Kuldanek S.A.,  
RA Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 4-615 FROM N.A.  
RX MEDLINE=86176794; PubMed=3754331;  
RA Tripodi M., Citarella F., Guida S., Galeffi P., Fantoni A.,  
RA Cortese R.;  
RT "cDNA sequence coding for human coagulation factor XII (Hageman).";  
RL Nucleic Acids Res. 14:3146-3146(1986).  
RN [4]  
RP SEQUENCE OF 14-615 FROM N.A.  
RX MEDLINE=86033830; PubMed=3877053;  
RA Cool D.E., Edgell C.-J.S., Louie G.V., Zoller M.J., Brayer G.D.,  
RA McGillivray R.T.A.;  
RT "Characterization of human blood coagulation factor XII cDNA.  
RT Prediction of the primary structure of factor XII and the tertiary  
RT structure of beta-factor XIIa.";  
RL J. Biol. Chem. 260:13666-13676(1985).  
RN [5]  
RP SEQUENCE OF 146-615 FROM N.A.  
RX MEDLINE=86216049; PubMed=3011063;  
RA Que B.G., Davie E.W.;  
RT "Characterization of a cDNA coding for human factor XII (Hageman  
RT factor).";  
RL Biochemistry 25:1525-1528(1986).  
RN [6]  
RP SEQUENCE OF 20-379.  
RX MEDLINE=85182674; PubMed=3886654;  
RA McMullen B.A., Fujikawa K.;  
RT "Amino acid sequence of the heavy chain of human alpha-factor XIIa  
RT (activated Hageman factor).";  
RL J. Biol. Chem. 260:5328-5341(1985).  
RN [7]  
RP SEQUENCE OF 354-362 AND 373-615.  
RX MEDLINE=83291041; PubMed=6604055;  
RA Fujikawa K., McMullen B.A.;  
RT "Amino acid sequence of human beta-factor XIIa.";  
RL J. Biol. Chem. 258:10944-10953(1983).  
RN [8]  
RP SEQUENCE OF 561-615 FROM N.A.  
RC TISSUE=Blood;

RX MEDLINE=96133302; PubMed=8528215;  
RA Schloesser M., Hofferbert S., Bartz U., Lutze G., Lammle B., Engel W.;  
RT "The novel acceptor splice site mutation 11396(G-->A) in the factor  
RT XII gene causes a truncated transcript in cross-reacting material  
RT negative patients.";  
RL Hum. Mol. Genet. 4:1235-1237(1995).  
RN [9]  
RP CARBOHYDRATE-LINKAGE SITE THR-109.  
RX MEDLINE=92184750; PubMed=1544894;  
RA Harris R.J., Ling V.T., Spellman M.W.;  
RT "O-linked fucose is present in the first epidermal growth factor  
RT domain of factor XII but not protein C.";  
RL J. Biol. Chem. 267:5102-5107(1992).  
RN [10]  
RP VARIANT WASHINGTON D.C. SER-590.  
RX MEDLINE=90046788; PubMed=2510163;  
RA Miyata T., Kawabata S.-I., Iwanaga S., Takahashi I., Alving B.,  
RA Saito H.;  
RT "Coagulation factor XII (Hageman factor) Washington D.C.: inactive  
RT factor XIIa results from Cys-571-->Ser substitution.";  
RL Proc. Natl. Acad. Sci. U.S.A. 86:8319-8322(1989).  
RN [11]  
RP VARIANT LOCARNO PRO-372.  
RX MEDLINE=94325559; PubMed=8049433;  
RA Hovington J.K., Schaller J., Stricker H., Willemin W.A., Furlan M.,  
RA Lammle B.;  
RT "Coagulation factor XII Locarno: the functional defect is caused by  
RT the amino acid substitution Arg-353-->Pro leading to loss of a  
RT kallikrein cleavage site.";  
RL Blood 84:1173-1181(1994).  
RN [12]  
RP VARIANT TENRI CYS-53.  
RX MEDLINE=99290785; PubMed=10361128;  
RA Kondo S., Tokunaga F., Kawano S., Ono Y., Kumagai S., Koide T.;  
RT "Factor XII Tenri, a novel cross-reacting material negative factor XII  
RT deficiency, occurs through a proteasome-mediated degradation.";  
RL Blood 93:4300-4308(1999).  
CC -I- FUNCTION: Factor XII is a serum glycoprotein that participates in  
CC the initiation of blood coagulation, fibrinolysis, and the  
CC generation of bradykinin and angiotensin.  
CC -I- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Ile bonds in factor  
CC VII to form factor VIIa and factor XI to form factor XIIa.  
CC -I- PTM: O- AND N-GLYCOSYLATED.  
CC -I- DISEASE: Defects in F12 do not cause any clinical symptoms. The  
CC sole effect is that whole-blood clotting time is prolonged.  
CC -I- MISCELLANEOUS: Factor XII, prekallikrein, and HMW kininogen form a  
CC complex bound to an anionic surface. Prekallikrein is cleaved by  
CC factor XII to form kallikrein, which then cleaves factor XII first  
CC to alpha-factor XIIa and then to beta-factor XIIa. Alpha-factor  
CC XIIa activates factor XI to factor Xla.  
CC -I- SIMILARITY: Belongs to peptidase family S1.  
CC -I- SIMILARITY: Contains 2 EGF-like domains.  
CC -I- SIMILARITY: Contains 1 fibronectin type I domain.  
CC -I- SIMILARITY: Contains 1 fibronectin type II domain.  
CC -I- SIMILARITY: Contains 1 kringle domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collabora-  
CC tion between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercia-  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M31315; AAA70225.1; -.  
CC EMBL; AF538691; AAM97932.1; -.  
CC EMBL; M11723; AAA51986.1; -.  
CC EMBL; M17465; AAB59490.1; -.  
CC EMBL; M17464; AAB59490.1; JOINED.  
CC EMBL; M17465; AAB59490.1; JOINED.  
CC EMBL; M13147; AAA70224.1; -.  
CC EMBL; U71274; AAB51203.1; -.  
CC PIR; A29411; KFHU12.

DR HSP; P00763; LDPO.  
DR MEROPS; S01.211; -  
DR Genew; HGNC:3530; F12.  
DR MIM; 234000; -  
DR GO; GO:0003805; F.blood coagulation factor XI activity; TAS.  
DR GO; GO:0003806; F.blood coagulation factor XII activity; TAS.  
DR GO; GO:0008236; F.serine-type peptidase activity; TAS.  
DR GO; GO:0007596; P.blood coagulation; TAS.  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR000083; Fibrinctn1.  
DR InterPro; IPR000562; FN\_Type\_II.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00008; EGF\_2.  
DR Pfam; PF00039; fn1; 1.  
DR Pfam; PF00040; fn2; 1.  
DR Pfam; PF00051; kringle; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00013; ENTPEPIL.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000995; FN\_Type\_II; 1.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00181; EGF\_2.  
DR SMART; SM00058; FN1; 1.  
DR SMART; SM00059; FN2; 1.  
DR SMART; SM00130; KR\_1.  
DR SMART; SM00020; Iryp\_Spc; 1.  
DR PROSITE; PS00022; EGF\_1; 2.  
DR PROSITE; PS01186; EGF\_2; 1.  
DR PROSITE; PS00026; EGF\_3; 2.  
DR PROSITE; PS01253; FIBRONECTIN\_1; 1.  
DR PROSITE; PS00023; FIBRONECTIN\_2; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS00070; KRINGLE\_2; 1.  
DR PROSITE; PS00400; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Glycoprotein; Blood coagulation; plasma; Kringle; Serine protease;  
KW Hydrolase; Fibrinolysis; Signal; EGF-like domain; Repeat; Zymogen;  
KW Polymorphism; Disease mutation.  
FT SIGNAL 1 39  
FT CHAIN 20 372 ALPHA-FACTOR XIIA HEAVY CHAIN.  
FT CHAIN 373 615 ALPHA-FACTOR XIIA LIGHT CHAIN.  
FT CHAIN 354 362 BETA-FACTOR XIIA PART 1.  
FT CHAIN 373 615 BETA-FACTOR XIIA PART 2.  
FT DOMAIN 47 88 FIBRONECTIN TYPE-II.  
FT DOMAIN 94 131 EGF-LIKE 1.  
FT DOMAIN 133 173 FIBRONECTIN TYPE-I.  
FT DOMAIN 174 210 EGF-LIKE 2.  
FT DOMAIN 217 295 KRINGLE.  
FT DOMAIN 296 349 PRO-RICH.  
FT DOMAIN 373 615 SERINE PROTEASE.  
FT CARBOHYD 109 109 O-LINKED (FUC).  
FT CARBOHYD 249 249 N-LINKED (GLCNAC...).  
FT CARBOHYD 299 299 O-LINKED (POTENTIAL).  
FT CARBOHYD 305 305 O-LINKED (POTENTIAL).  
FT CARBOHYD 308 308 O-LINKED (POTENTIAL).  
FT CARBOHYD 328 328 O-LINKED (POTENTIAL).  
FT CARBOHYD 329 329 O-LINKED (POTENTIAL).  
Query Match 33.4%; Score 170.5; DB 1; Length 615;  
Best Local Similarity 48.5%; Pred.No.58-12; 24; Indels 5; Gaps 2;  
Matches 33; Conservative 6; Mismatches 24; Indels 5; Gaps 2;  
QY 2 TYENGNGHYRGKASDTWGRCPLPNSATVLCQTY-HAHRSDALQGLGKHNYCRNPDN 60  
Db 216 SCVDGRGLSYRGARTLTLGAPCPQWAS----EATYRVNTAEQARNWGLGHAFCRFDN 271  
QY 61 RRRPWCYV 68

Db 272 DIRPWCYV 279  
RESULT 20  
APOA\_MACMU  
ID APOA\_MACMU STANDARD; PRT; 1420 AA.  
AC P14417;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Apolipoprotein(a) (EC 3.4.21.-) (Apo(a)) (Lp(a)) (Fragment).  
GN LpA.  
OS Macaca mulatta (Rhesus macaque).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecoidea; Macaca.  
OX NCBI\_TaxID=9544;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89174660; PubMed=29256643;  
RA Tomlinson J.E., McLean J.W., Lawn R.M.;  
RT "Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of  
synthesis";  
RL J. Biol. Chem. 264:5957-5965(1989).  
CC -!- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)  
(Lp(a)). It has serine proteinase activity and is able of  
autoproteolysis. Inhibits tissue-type plasminogen activator 1.  
CC Lp(a) may be a ligand for megalin/Cp330.  
CC -!- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and  
decorin (By similarity).  
CC -!- PTM: N- and O-glycosylated (By similarity).  
CC -!- DISEASE: Elevated plasma concentrations of apo(a) and its  
naturally occurring proteolytic fragments are correlated with  
atherosclerosis. Homology with plasminogen kringle IV and V is  
thought to underlie the atherogenicity of the protein, because the  
fragments are competing with plasminogen for fibrinogen binding.  
CC -!- MISCELLANEOUS: Apo(a) is known to be proteolytically cleaved,  
leading to the formation of the so called mini-Lp(a). Apo(a)  
fragments accumulate in atherosclerotic lesions, where they may  
promote thrombogenesis. O-glycosylation may limit the extent of  
proteolytic fragmentation (By similarity).  
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.  
CC -!- SIMILARITY: Contains at least 10 kringle domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
at the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; J04635; AAA36833.1; -.  
CC PIR; A32869; A32869.  
CC HSP; P00747; 2PK4.  
CC MEROPS; S01.226;  
CC InterPro; IPR009003; Cys\_Ser\_trypsin.  
CC InterPro; IPR000001; Kringle.  
CC InterPro; IPR001254; Peptidase\_S1.  
CC InterPro; IPR001314; Peptidase\_S1A.  
CC Pfam; PF00051; kringle; 1.  
CC Pfam; PF00089; trypsin; 1.  
CC PRINTS; PR00722; CHYMOTRYPSIN.  
CC PRINTS; PR00018; KRINGLE.  
CC ProDom; PD000395; Kringle; 10.  
CC SMART; SM00130; KR; 10.  
CC SMART; SM00020; TRYPSIN\_SPC; 1.  
CC PROSITE; PS00021; KRINGLE\_1; 10.  
CC PROSITE; PS00070; KRINGLE\_2; 10.  
CC PROSITE; PS00240; TRYPSIN\_DOM; 1.  
CC PROSITE; PS00134; TRYPSIN\_HIS; FALSE\_NEG.  
CC PROSITE; PS00135; TRYPSIN\_SER; FALSE\_NEG.

Hydrolase; Serine protease; Lipid transport; Plasma; Glycoprotein;  
 KW Kringle; Repeat; Atherosclerosis.  
 FT NON TER 1 127 KRINGLE 1.  
 FT DOMAIN 49 163 KRINGLE 2.  
 FT DOMAIN 163 241 KRINGLE 3.  
 FT DOMAIN 277 355 KRINGLE 4.  
 FT DOMAIN 391 469 KRINGLE 5.  
 FT DOMAIN 505 583 KRINGLE 6.  
 FT DOMAIN 619 697 KRINGLE 7.  
 FT DOMAIN 725 803 KRINGLE 8.  
 FT DOMAIN 839 917 KRINGLE 9.  
 FT DOMAIN 953 1031 KRINGLE 10.  
 FT DOMAIN 1067 1145 SERINE PROTEASE.  
 FT DOMAIN 1191 1420  
 SQ SEQUENCE 1420 AA; 158367 MW; BE102949E03C5B0E CRC64;  
 Query Match 31.6%; Score 161; DB 1; Length 1420;  
 Best Local Similarity 40.9%; Pred. No. 15e-10;  
 Matches 36; Conservative 9; Mismatches 31; Indels 12; Gaps 4;  
 QY 3 CYEGNGHYRGKASTDTMGPRCPWNSATVLQ--QTYHAHRSDALQGLGKHNYCRNPDN 60  
 Db 1068 CYHNGQSYRGFTSTTGTGTCQSSWMTPHQHKRTPEHNPDDLTM-----NYCRNPDA 1122  
 QY 61 RRRPWCYVQGLKPLVQE--CMVHDCAD 86  
 Db 1123 DTGPWCFT--MDPSVREYCNLTRCSD 1147  
 RESULT 21  
 KRM1\_MOUSE  
 ID KRM1\_MOUSE STANDARD; PRT; 473 AA.  
 AC Q99N43;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Kremen protein 1 precursor (Kringle-containing protein marking the eye  
 and the nose) (Dickkopf receptor).  
 GN KREMEN1 OR KREMEN.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., DEVELOPMENTAL STAGE, AND TISSUE SPECIFICITY.  
 RC TISSUE=Brain, and Kidney;  
 RX MEDLINE=21167372; PubMed=11267660;  
 RA Nakamura T., Aoki S., Kitajima K., Takahashi T., Matsumoto K.,  
 RA Nakamura T.;  
 RT "Molecular cloning and characterization of Kremen, a novel  
 RT kringle-containing transmembrane protein.";  
 RL Biochim. Biophys. Acta 1518:63-72(2001).  
 CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf  
 CC to block Wnt/beta-catenin signaling (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).  
 CC -!- TISSUE SPECIFICITY: In the adult, widely expressed with high  
 CC levels in heart, lung, kidney, skeletal muscle and testis.  
 CC -!- DEVELOPMENTAL STAGE: In the embryo, expression is first detected  
 CC on day 9 and increases up to day 18. Lower levels are found in  
 CC adult. At 9.5 dpc, expression is localised to the apical  
 CC ectodermal ridge (AER) of the developing fore- and hindlimb buds,  
 CC the telencephalon and the first brachial arch. At 10.5 dpc,  
 CC expression is also observed in the myotome and in sensory tissues  
 CC such as the nasal pit and optic vesicle.  
 CC -!- SIMILARITY: Contains 1 CUB domain.  
 CC -!- SIMILARITY: Contains 1 Kringle domain.  
 CC -!- SIMILARITY: Contains 1 WSC domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce.  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; AB059517; BAB40968.1; -.  
 DR HSSP; P00747; ICEA.  
 DR MGD; MGI:1933988; Kremen.  
 DR GO; GO:0016021; C:integral to membrane; NAS.  
 DR InterPro; IPR000859; CUB.  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR002889; WSC.  
 DR Pfam; PF00431; CUB; 1.  
 DR Pfam; PF00051; Kringle; 1.  
 DR Pfam; PF01822; WSC; 1.  
 DR PRINTS; PRO0018; KRINGLE.  
 DR ProDom; PD000395; Kringle; 1.  
 DR SMART; SM00042; CUB; 1.  
 DR SMART; SM00130; KR; 1.  
 DR PROSITE; PS01180; CUB; 1.  
 DR PROSITE; PS00021; KRINGLE\_1; 1.  
 DR PROSITE; PS00070; KRINGLE\_2; 1.  
 KW Wnt signaling pathway; Signal; Transmembrane; Kringle.  
 FT SIGNAL 1 19  
 FT CHAIN 20 473 KREMEN PROTEIN 1.  
 FT DOMAIN 21 392 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 393 413 POTENTIAL.  
 FT DOMAIN 414 473 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 31 114 KRINGLE.  
 FT DOMAIN 120 210 WSC.  
 FT DOMAIN 214 321 CUB.  
 FT CARBOHYD 59 59 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 217 217 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 255 255 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 293 293 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 333 333 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 345 345 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 473 AA; 51716 MW; 58682778BE3FDD1 CRC64;  
 Query Match 31.2%; Score 159; DB 1; Length 473;  
 Best Local Similarity 45.6%; Pred. No. 8.1e-11;  
 Matches 31; Conservative 7; Mismatches 26; Indels 4; Gaps 2  
 QY 3 CYEGNGHYRGKASTDTM--GRCPWPNSATVLQTYHAHRSDALQGLGKHNYCRNPDN 60  
 Db 32 CFTANGADYRGTSQWALQGGKPCLFWNE--TFQHPYNTLKYPNVGGGLGKHNYCRNPDG 89  
 QY 61 RRRPWCYV 68  
 Db 90 DVSPWCYV 97  
 RESULT 22  
 KRM1\_RAT  
 ID KRM1\_RAT STANDARD; PRT; 473 AA.  
 AC Q92454;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Kremen protein 1 precursor (Kringle-containing protein marking the eye  
 and the nose) (Dickkopf receptor).  
 GN KREMEN1 OR KREMEN.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Nakamura T., Nakamura T.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf  
 CC to block Wnt/beta-catenin signaling (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).  
 CC -!- SIMILARITY: Contains 1 CUB domain.  
 CC -!- SIMILARITY: Contains 1 kringle domain.

--!- SIMILARITY: Contains 1 WSC domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

EMBL; AB065090; BAB62003.1; -.
GO; GO:0016021; C:integral to membrane; ISS.
InterPro; IPR000859; CUB.
InterPro; IPR000001; Kringle.
InterPro; IPR002889; WSC.
Pfam; PF00431; CUB; 1.
Pfam; PF00051; kringle; 1.
Pfam; PF01822; WSC; 1.
PRINTS; PR00018; KRINGLE.
ProDom; PD000395; Kringle; 1.
SMART; SM00042; CUB; 1.
SMART; SM00130; KR; 1.
PROSITE; PS001180; CUB; 1.
PROSITE; PS00021; KRINGLE_1; 1.
PROSITE; PS50070; KRINGLE_2; 1.
Wnt signaling pathway; Signal; Transmembrane; Kringle.

```

FT	1	19	POTENTIAL.
FT	20	473	KREMEN PROTEIN 1.
FT	21	392	EXTRACELLULAR (POTENTIAL).
FT	393	413	POTENTIAL.
FT	414	473	CYTOPLASMIC (POTENTIAL).
FT	31	114	KRINGLE.
FT	120	210	WSC.
FT	214	321	CUB.
FT	59	59	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	217	217	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	255	255	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	293	293	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	333	333	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	345	345	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	473 AA:	51869 MW:	9B510857DF0856F08 CRC64:
SQ	SEQUENCE		

```

Query Match      31.2%; Score 159; DB 1; Length 473;
Best Local Similarity 45.8%; Pred. No. 8.1e-11;
Matches 31; Conservative 7; Mismatches 26; Indels 4; Gaps 2;

QY 3 CVEGNGHYRGKASTDTM--GRPCLPWNSATVLLQOTYHAHRSDALQGLGKHNYCRNPDN 60
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 32 CFTAGADYRGTSWTALQGGKPLFWE--TFQHFYNTLKYPNGEGLGEHNYCRNPDG 89
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 61 RRRPWCYV 68
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 90 DVSPPWCYV 97
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 23	
KRMI_XENLA	
ID	KRMI_XENLA STANDARD; PRT; 452 AA.
AC	Q90Y30;
DT	10-OCT-2003 (Rel. 42, Created)
DT	10-OCT-2003 (Rel. 42, Last sequence update)
DT	10-OCT-2003 (Rel. 42, Last annotation update)
DE	Kremen protein 1 precursor [Krigle-containing protein marking the eye and the nose] (Dickkopf receptor).
DE	
GN	KREMEN1
OS	Xenopus laevis (African clawed frog).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC	Xenopodinae; Xenopus.
OX	NCBI_TaxID=8355;
[1]	
RN	
RP	SEQUENCE FROM N.A.
RA	Nakamura T.

```

Submitted (AUG-2001) to the EMBL/GenBank/DDJF databases.
-!- FUNCTION: Receptor for dickkopf protein. Cooperates with Dickkopf
      to block Wnt/beta-catenin signaling (By similarity).
-!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
-!- SIMILARITY: Contains 1 CUB domain.
-!- SIMILARITY: Contains 1 kringle domain.
-!- SIMILARITY: Contains 1 WSC domain.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce
or send an email to license@isb-sib.ch).

```

DR	EMBL; AB070851; BA364294.1; .	
DR	InterPro; IPR000859; CUB.	
DR	InterPro; IPR000001; Kringle.	
DR	InterPro; IPR002889; WSC.	
DR	Pfam; PF00431; CUB; 1.	
DR	Pfam; PF00051; kringle; 1.	
DR	Pfam; PF01822; WSC; 1.	
DR	PRINTS; PR00018; KRINGLE.	
DR	ProDom; PD000395; Kringle; 1.	
DR	SMART; SMO0042; CUB; 1.	
DR	SMART; SMO0130; KR; 1.	
DR	SMART; SMO0321; WSC; 1.	
DR	PROSITE; PS01180; CUB; 1.	
DR	PROSITE; PS00021; KRINGLE; 1; 1.	
DR	PROSITE; PS00070; KRINGLE2; 1.	
DR	Wnt signaling pathway; Glycoprotein; Kringle;	
KW	transmembrane; Signal.	
FT	SIGNAL 1 22	POTENTIAL.
FT	CHAIN 23 452	KREMEN PROTEIN 1.
FT	DOMAIN 23 369	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 370 390	POTENTIAL.
FT	DOMAIN 391 452	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN 29 112	KRINGLE.
FT	DOMAIN 118 208	WSC.
FT	DOMAIN 212 319	CUB.
FT	CARBOHYD 43 43	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 57 57	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 215 215	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 253 253	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 291 291	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 328 328	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 344 344	N-LINKED (GLCNAC. .) (POTENTIAL).
DR	SEQUENCE 452 AA; 50188 MW; BD24BCD1AF4564E2 CRC64;	

Query Match	31.0%;	Score 158;	DB 1;	Length 452;
Best Local Similarity	44.1%;	Pred. No. 1e-10;		
Matches 30:	Conservative	9;	Mismatches 25;	Indels 4;
	Gap	2;		

	QY	3	CYEANGHYFGKASTDTM--GRCLPWN SATVLQQTYHAHRSDALQLGLGKNYCRNPDN	60
Dd		30	:       : :       :	
	Dd		CYTNGADYRGTONQTSLDGGKCLFWE--TFQH FYNTLKYPNEGEGLGEHNYCRNPDG	87
QY		61	RRRPWCYV	68
Dd		88	DVSPWCVI	95

RESULT 24	
ID	KRM1 HUMAN
ID	KRM1 HUMAN
STANDARD;	PRT; 475 AA.
Q96MÜ8; Q9UB70; Q9UGS5; Q9UGU1;	
28-FEB-2003 (Rel. 41, Created)	
DT	
28-FEB-2003 (Rel. 41, Last sequence update)	
DT	
10-OCT-2003 (Rel. 42, Last annotation update)	
DE	
DE	Kremen protein 1 precursor (Krigle-containing protein marking the eye
DE	and the nose) (ickkopf receptor).
GN	KREMEN1 OR KREMEN.

OS Homo sapiens (Human)  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RA Nakamura T., Nakamura T.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,  
 RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,  
 RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,  
 RA Matsui H., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,  
 RA Murakawa K., Kamehori K., Takahashi-Fujii K., Oshima K., Sugiyama A.,  
 RA Kawakami B., Suzuki Y., Sugano S., Nagahara K., Masuho Y., Nagai K.,  
 RA Isogai T.;  
 RL "NEDO human cDNA sequencing project";  
 RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 EX MEDLINE=20057165; PubMed=10591208;  
 RA Dunham I., Hunt A.R., Collins J.E., Bruskiewich R., Beare D.M.,  
 RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,  
 RA Baguley C., Bailey J., Ballow K.F., Bates K.N., Beasley O.P.,  
 RA Bird C.P., Blakey S.E., Bridgman A.M., Buck D., Burgess J.,  
 RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,  
 RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,  
 RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,  
 RA Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,  
 RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,  
 RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,  
 RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,  
 RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,  
 RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,  
 RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., McCann O.T.,  
 RA McElay C.N., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,  
 RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,  
 RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,  
 RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,  
 RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,  
 RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,  
 RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,  
 RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,  
 RA Shintani A., Shibuya K., Yoshizaki Y., Acki N., Mitsuyama S.,  
 RA Roe B.A., Chen F., Chu L., Crabtree J., Acki N., Mitsuyama S.,  
 RA Lewis J., Lewis S., Lin S.-P., Loh P., Malj E., Nguyen T., Pan H.,  
 RA Phan S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,  
 RA Zhan Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,  
 RA Fulton R., Johnson D., Bemis G., Murray J., Miller N., Minx P.,  
 RA Cordes M., Du Z., Fulton L., Goela D., Bradshaw H., Bourne S.,  
 RA Hinds K., Kemp K., Latreille P., Layman D., Ozeresky P., Rohlfing T.,  
 RA Scheet P., Walker C., Wamley A., Wohlmann P., Pepin K., Nelson J.,  
 RA Korf I., Bedell J.A., Hillier L.W., Mardis E., Waterston R.,  
 RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,  
 RA Edlmann L., Kim U.J., Shizuya H., Simon M.I., Dumanowski J.P.,  
 RA Peyrard M., Kedra D., Seroussi E., Franson I., Tapia I., Bruder C.E.,  
 RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,  
 RA Khan A.S., Lane L., Tilahun Y., Wright H.;  
 RT "The DNA sequences of human chromosome 22";  
 RL Nature 402:489-495(1999).  
 CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf  
 CC to block Wnt/beta-catenin signaling (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=Q96MU8-1; Sequence=Displayed;

Name=2;  
 IsoId=Q96MU8-2; Sequence=VSP 003900;  
 Notes=No experimental confirmation available;  
 -!- SIMILARITY: Contains 1 CUB domain.  
 -!- SIMILARITY: Contains 1 kringle domain.  
 -!- SIMILARITY: Contains 1 WSC domain.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collabora-  
 CC tion between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercia  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce  
 CC or send an email to license@isb-sib.ch).  
 -----  
 CC EMBL; AB059618; BAB40969.1; -;  
 CC EMBL; AK056425; BAB71180.1; -;  
 CC EMBL; Z95116; CAB62952.1; -;  
 CC EMBL; AL021393; CAB62959.1; -;  
 CC Genew; HGNC:17550; K3EMEN1;  
 CC GO; GO:0016021; C:integral to membrane; ISS.  
 CC GO; GO:0005624; C:membrane fraction; TAS.  
 CC GO; GO:0007154; P:cell communication; TAS.  
 CC InterPro; IPR000859; CUB.  
 CC InterPro; IPR000001; Kringle.  
 CC InterPro; IPR002889; WSC.  
 CC Pfam; PF00431; CUB; 1.  
 CC Pfam; PF00051; kringle; 1.  
 CC Pfam; PF01822; WSC; 1.  
 CC PRINTS; PR00018; KRINGLE.  
 CC PRODOM; PD000395; Kringle; 1.  
 CC SMART; SM00042; CUB; 1.  
 CC SMART; SM00130; KR; 1.  
 CC PROSITE; PS01180; CUB; 1.  
 CC PROSITE; PS00021; KRINGLE; 1; 1.  
 CC PROSITE; PS00070; KRINGLE\_2; 1.  
 CC Wnt signaling pathway; Signal; Transmembrane; Kringle;  
 CC Alternative splicing.  
 FT SIGNAL 1 19  
 FT CHAIN 20 475  
 FT DOMAIN 21 394  
 FT TRANSMEM 395 415  
 FT DOMAIN 416 475  
 FT DOMAIN 121 202  
 FT DOMAIN 216 323  
 FT CARBOHYD 47 47  
 FT CARBOHYD 61 61  
 FT CARBOHYD 219 219  
 FT CARBOHYD 295 295  
 FT CARBOHYD 335 335  
 FT CARBOHYD 347 347  
 FT VARSPPLIC 473 475  
 FT FT  
 FT FTID=VSP 003900.  
 FT CONFLICT 29 30 MISSING (IN REF. 1).  
 FT CONFLICT 206 206 I -> V (IN REF. 2).  
 FT SEQUENCE 475 AA; 51898 MW; B7B86F080F96A04 CRC64;  
 -----  
 Query Match 30.6%; Score 156; DB 1; Length 475;  
 Best Local Similarity 44.1%; Pred. No. 1.8e-10;  
 Matches 30; Conservative 8; Mismatches 26; Indels 4; Gaps 2  
 OY 3 CTEGNGHFYRGKASTDTM--GRPCLPWNSATVLTQTYHAHRSDALQLGLKKNYCRNDP 60  
 DB 34 CFTANGADYRGTCQNTALQGGKPCLFWNS--TFQHPYNTLKYPNGGGLGEHNYCRNDPG 91  
 OY 61 RRRPWCYV 68  
 DB 92 DVSFPCYV 99  
 -----  
 RESULT 25



APOA HUMAN  
ID APOA HUMAN STANDARD; PRT; 4548 AA.  
AC P08519;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 01-AUG-1988 (Rel. 08, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Apolipoprotein(a) precursor (EC 3.4.21.-) (Apo(a)) (Lp(a)).  
GN LPA.

OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1] SEQUENCE FROM N.A.  
RX MEDLINE=88039109; PubMed=3670400;  
RA McLean J.W., Tomlison J.E., Kuang W.-J., Eaton D.L., Chen E.Y.,  
RA Fless G.M., Scam A.M., Lawn R.M.;  
RT "cDNA sequence of human apolipoprotein(a) is homologous to  
RT plasminogen.";  
RL Nature 330:132-137(1987).  
RN [2]  
RP SERINE PROTEASE ACTIVITY.  
RX MEDLINE=90076123; PubMed=2531657;  
RA Salonen E.-M., Jauhainen M., Zardi L., Vaheri A., Enholm C.;  
RT "Lipoprotein(a) binds to fibronectin and has serine proteinase  
RT activity capable of cleaving it.";  
RL EMBO J. 8:4035-4040(1989).  
RN [3]  
RP REVIEW.  
RX MEDLINE=90049223; PubMed=2330631;  
RA Utermann G.;  
RT "The mysteries of lipoprotein(a).";  
RL Science 246:904-910(1989).  
RN [4]  
RP CHARACTERIZATION OF THE N- AND O-LINKED GLYCANS.  
RX MEDLINE=21103595; PubMed=11294842;  
RA Garner B., Merry A.H., Royle L., Harvey D.J., Rudd P.M., Thillet J.;  
RT "Structural elucidation of the N- and O-glycans of human  
RT apolipoprotein(a): role of o-glycans in conferring protease  
RT resistance.";  
RL J. Biol. Chem. 276:22200-22208(2001).  
RN [5]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 4121-4208.  
RX MEDLINE=96217891; PubMed=8642595;  
RA Mikol V., Lograsso P.V., Beetcher B.R.;  
RT "Crystal structures of apolipoprotein(a) kringle IV37 free and  
RT complexed with 6-aminohexanoic acid and with p-aminomethylbenzoic  
RT acid: existence of novel and expected binding modes.";  
RL J. Mol. Biol. 256:751-761(1996).  
RN [6]  
RP VARIANT ARG-4193.  
RX MEDLINE=95002201; PubMed=7918692;  
RA Scam A.M., Pfaffinger D., Lee J.C., Hinman J.;  
RT "A single point mutation (Trp72-->Arg) in human apo(a) kringle 4-37  
RT associated with a lysine binding defect in Lp(a).";  
RL Biochim. Biophys. Acta 1227:41-45(1994).  
CC -!- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)  
CC (Lp(a)). It has serine proteinase activity and is able of  
CC autolysis. Inhibits tissue-type plasminogen activator 1.  
CC Lp(a) may be a ligand for megalin/Gp 330.  
CC -!- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and  
CC decorin.  
CC -!- PTM: N- and O-glycosylated. The N-glycans are complex biantennary  
CC structures present in either a mono- or disialylated state. The  
CC O-glycans are mostly (80%) represented by the monosialylated core  
CC type I structure, NeuNacalpa2-3Galbeta1-3GalNAc, with smaller  
CC amounts of disialylated and non-sialylated O-glycans also  
CC detected.  
CC -!- DISEASE: Elevated plasma concentrations of apo(a) and its  
CC naturally occurring proteolytic fragments are correlated with  
CC atherosclerosis. Homology with plasminogen kringle IV and V is  
CC thought to underlie the atherogenicity of the protein, because the  
CC fragments are competing with plasminogen for fibrin(ogen) binding.

-!- MISCELLANEOUS: Apo(a) is known to be proteolytically cleaved,  
leading to the formation of the so called mini-Lp(a). Apo(a)  
fragments accumulate in atherosclerotic lesions, where they may  
promote thrombogenesis. O-glycosylation may limit the extent of  
proteolytic fragmentation.  
-!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.  
-!- SIMILARITY: Contains 38 kringle domains.  
This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation  
at the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announcement>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
EMBL; X06290; CAA29618.1; -  
PIR; S00657; S00657.  
PDB; 1J71; 13-JUN-01.  
PDB; 1JFN; 28-JUN-02.  
PDB; 1KIV; 18-MAY-99.  
PDB; 3KIV; 18-MAY-99.  
PDB; 4KIV; 18-MAY-99.  
MEROPS; S01.226; -  
Genew; HGNC:6667; LPA.  
MTM; 152200; -  
GO; GO:0004866; F:endorpeptidase inhibitor activity; TAS.  
GO; GO:0008015; P:circulation; TAS.  
GO; GO:0009405; P:pathogenesis; TAS.  
InterPro; IPR009003; Cys Ser trypsin.  
InterPro; IPR000001; Kringle.  
InterPro; IPR001254; Peptidase S1.  
InterPro; IPR001314; Peptidase\_S1A.  
Pfam; PF00051; kringle; 38.  
Pfam; PF00089; trypsin; 1.  
PRINTS; PR00722; CHYMOTRYPSIN.  
PRINTS; PR00018; KRINGLE.  
ProDom; PD000395; Kringle; 38.  
SMART; SM00130; KR; 38.  
SMART; SM00020; Tryp\_SPC; 1.  
PROSITE; PS00021; KRINGLE\_1; 38.  
PROSITE; PS00070; KRINGLE\_2; 38.  
PROSITE; PS00240; TRYPSIN\_DOM; 1.  
PROSITE; PS00134; TRYPSIN\_HIS; 1.  
PROSITE; PS00135; TRYPSIN\_SER; 1.  
Hydrolase; Serine protease; Lipid transport; Plasma; Glycoprotein;  
Kringle; Repeat; Atherosclerosis; Signal; Polymorphism; 3D-structure.  
FT SIGNAL 1 19  
FT CHAIN 20 4548 APOLIPOPROTEIN(A).  
FT DOMAIN 20 130 KRINGLE TYPE IV, 1.  
FT DOMAIN 131 244 KRINGLE TYPE IV, 2.  
FT DOMAIN 245 358 KRINGLE TYPE IV, 3.  
FT DOMAIN 359 472 KRINGLE TYPE IV, 4.  
FT DOMAIN 473 586 KRINGLE TYPE IV, 5.  
FT DOMAIN 587 700 KRINGLE TYPE IV, 6.  
FT DOMAIN 701 814 KRINGLE TYPE IV, 7.  
FT DOMAIN 815 928 KRINGLE TYPE IV, 8.  
FT DOMAIN 929 1042 KRINGLE TYPE IV, 9.  
FT DOMAIN 1043 1156 KRINGLE TYPE IV, 10.  
FT DOMAIN 1157 1270 KRINGLE TYPE IV, 11.  
FT DOMAIN 1271 1384 KRINGLE TYPE IV, 12.  
FT DOMAIN 1385 1498 KRINGLE TYPE IV, 13.  
FT DOMAIN 1499 1612 KRINGLE TYPE IV, 14.  
FT DOMAIN 1613 1726 KRINGLE TYPE IV, 15.  
FT DOMAIN 1727 1840 KRINGLE TYPE IV, 16.  
FT DOMAIN 1841 1954 KRINGLE TYPE IV, 17.  
FT DOMAIN 1955 2068 KRINGLE TYPE IV, 18.  
FT DOMAIN 2069 2182 KRINGLE TYPE IV, 19.  
FT DOMAIN 2183 2296 KRINGLE TYPE IV, 20.  
FT DOMAIN 2297 2410 KRINGLE TYPE IV, 21.  
FT DOMAIN 2411 2524 KRINGLE TYPE IV, 22.  
FT DOMAIN 2525 2638 KRINGLE TYPE IV, 23.  
FT DOMAIN 2639 2752 KRINGLE TYPE IV, 24.

```
FT DOMAIN 2753 2866 KRINGLE TYPE IV, 25.
FT DOMAIN 2867 2980 KRINGLE TYPE IV, 26.
FT DOMAIN 2981 3094 KRINGLE TYPE IV, 27.
FT DOMAIN 3095 3208 KRINGLE TYPE IV, 28.
FT DOMAIN 3209 3322 KRINGLE TYPE IV, 29.
FT DOMAIN 3323 3436 KRINGLE TYPE IV, 30.
FT DOMAIN 3437 3550 KRINGLE TYPE IV, 31.
FT DOMAIN 3551 3664 KRINGLE TYPE IV, 32.
FT DOMAIN 3665 3770 KRINGLE TYPE IV, 33.
FT DOMAIN 3771 3884 KRINGLE TYPE IV, 34.
FT DOMAIN 3885 3998 KRINGLE TYPE IV, 35.
FT DOMAIN 3999 4112 KRINGLE TYPE IV, 36.
FT DOMAIN 4113 4226 KRINGLE TYPE IV, 37.
FT DOMAIN 4227 4340 KRINGLE TYPE V.
FT DOMAIN 4341 4454 SERINE PROTEASE.
FT ACT_SITE 4455 4568 CHARGE RELAY SYSTEM.
FT ACT_SITE 4569 4682 CHARGE RELAY SYSTEM.
FT ACT_SITE 4683 4796 CHARGE RELAY SYSTEM.
FT VARIANT 4797 4910 W -> R (LOSS OF LYSINE-SEPHAROSE BINDING).
FT SQ SEQUENCE 4548 AA; 501313 MM; 96921B596A465C5F CRC64;
      Query Match 30.6%; Score 156; DB 1; Length 4548;
      Best Local Similarity 37.8%; Pred. No. 1.9e-09;
      Matches 34; Conservative 12; Mismatches 32; Indels 12; Gaps 4;
QY 1 KTCYEGNGHFYRGKASTDMGRCPPLMNSATV-LQOTYHAHRSDALQLGLGKHYCRNP 58
Db 4122 RQCYHGNGSRYGTSTVTGRTCSWSMTPHRRQTPNPNDGLTM-----NYCRNP 4176
QY 59 DNRRRPWCYVQGLKPLV--QECMVHDCAD 86
Db 4177 DADTGPWCFT---MDPSIRWEYCNLTRGSD 4203

RESULT 26
ID FA12_BOVIN STANDARD; PRT; 593 AA.
AC P98140;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
DR (HAF) (Fragment).
GN F12.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94242782; PubMed=8186251;
RA Shibuya Y., Semba U., Okabe H., Kambara T., Yamamoto T.;
RT "Primary structure of bovine Hageman factor (blood coagulation factor
RT XII): comparison with human and guinea pig molecules."
RL Biochim. Biophys. Acta 1206:63-70(1994).
RP [2]
RP SEQUENCE OF 10-21; 350-364 AND 525-550.
RX MEDLINE=77182112; PubMed=861210;
RA Fujikawa K., Walsh A.K., Davie W.E.;
RT "Isolation and characterization of bovine factor XII (Hageman
RT factor).".
RL Biochemistry 16:2270-2278(1977).
CC -!- FUNCTION: Factor XII is a serum glycoprotein that participates in
CC the initiation of blood coagulation, fibrinolysis, and the
CC generation of bradykinin and angiotensin.
CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Ile bonds in factor
CC VII to form factor VIIa and factor XI to form factor XIa.
CC -!- PTM: O- AND N-GLYCOSYLATED (BY SIMILARITY).
CC -!- MISCELLANEOUS: Factor XII, prekallikrein, and HMW kininogen form a
```

```
CC complex bound to an anionic surface. Prekallikrein is cleaved by
CC factor XII to form kallikrein, which then cleaves factor XII first
CC to alpha-factor XIIa and then to beta-factor XIIa. Alpha-factor
CC XIIa activates factor XI to factor XIa. Bovine factor XII is
CC cleaved only to alpha-factor XIIa as it lacks the trypsin/
CC kallikrein cleavage site.
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboratio
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on it
CC use by non-profit institutions as long as its content is in no wa
CC modified and this statement is not removed. Usage by and for commercia
CC entities requires a license agreement (See http://www.isb-sib.ch/announce
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; S70164; AAB30804.2; --
CC PIR; S45281; S45281.
CC RSP; P00763; IDPO.
CC MEROPS; S01.211; -.
CC InterPro; IPR009003; Cys Ser trypsin.
CC InterPro; IPR006209; EGF-like.
CC InterPro; IPR000083; Fibrinctn1.
CC InterPro; IPR000562; FN_Type_II.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Peptidase_S1.
CC InterPro; IPR001314; Peptidase_S1A.
CC Pfam; PF00008; EGF_2.
CC Pfam; PF00039; fn1; 1.
CC Pfam; PF00040; fn2; 1.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00013; FNTYPEII.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000995; FN_Type_II; 1.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00181; EGF; 2.
CC SMART; SM00058; FN1; 1.
CC SMART; SM00059; FN2; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00020; Tryp_SPC; 1.
CC PROSITE; PS00022; EGF_1; 2.
CC PROSITE; PS01186; EGF_2; FALSE_NEG.
CC PROSITE; PS00026; EGF_3; 2.
CC PROSITE; PS01253; FIBRONECTIN_1; 1.
CC PROSITE; PS00023; FIBRONECTIN_2; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS00070; KRINGLE_2; 1.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
CC Hydrolyase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.
CC FT SIGNAL 1 9
CC FT CHAIN 10 349
CC FT CHAIN 350 593
CC FT DOMAIN 37 78
CC FT DOMAIN 84 121
CC FT DOMAIN 123 163
CC FT DOMAIN 164 200
CC FT DOMAIN 207 287
CC FT DOMAIN 297 333
CC FT DOMAIN 350 593
CC FT ACT_SITE 389 389
CC FT ACT_SITE 438 438
CC FT ACT_SITE 541 541
CC ALPHA-FACTOR XIIA HEAVY CHAIN.
CC ALPHA-FACTOR XIIA LIGHT CHAIN.
CC FIBRONECTIN TYPE-II.
CC EGF-LIKE 1.
CC EGF-LIKE 2.
CC FIBRONECTIN TYPE-I.
CC KRINGLE.
CC PRO-RICH.
CC SERINE PROTEASE.
CC CHARGE RELAY SYSTEM (BY SIMILARITY).
CC CHARGE RELAY SYSTEM (BY SIMILARITY).
CC CHARGE RELAY SYSTEM (BY SIMILARITY).
```

```

FT DISULFID 88 100 BY SIMILARITY.
FT DISULFID 94 109 BY SIMILARITY.
FT DISULFID 111 120 BY SIMILARITY.
FT DISULFID 125 153 BY SIMILARITY.
FT DISULFID 151 160 BY SIMILARITY.
FT DISULFID 168 179 BY SIMILARITY.
FT DISULFID 173 188 BY SIMILARITY.
FT DISULFID 190 199 BY SIMILARITY.
FT DISULFID 207 287 BY SIMILARITY.
FT DISULFID 230 269 BY SIMILARITY.
FT DISULFID 258 282 BY SIMILARITY.
FT DISULFID 336 463 BY SIMILARITY.
FT DISULFID 374 390 BY SIMILARITY.
FT DISULFID 392 452 BY SIMILARITY.
FT DISULFID 413 416 BY SIMILARITY.
FT DISULFID 479 547 BY SIMILARITY.
FT DISULFID 510 536 BY SIMILARITY.
FT DISULFID 537 568 BY SIMILARITY.
FT CARBOHYD 99 99 O-LINKED (FUC) (BY SIMILARITY).
FT CARBOHYD 241 241 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 263 263 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 410 410 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 593 AA; 65148 MW; 721592BA792BD61F CRC64;

Query Match 30.5%; Score 155.5; DB 1; Length 593;
Best Local Similarity 38.4%; Pred. No. 2.6e-10;
Matches 33; Conservative 9; Mismatches 37; Indels 7; Gaps 3;

QY 2 TCYE-GNCHFYRGKASTDTMGRPCLPWNSATVLOQTY-HAHRSDALGLGKKNYCRNP 58
Db 206 SCYDRDRSLYRGWAGTTLSPACQSWAS-----EATYNNVTAEQVLWGLGDAFCRNP 261

QY 59 DNRRRPWCYVQGLKPLVQECMVHDC 84
Db 262 DNDTRPWCYVQGLKPLVQECMVHDC 287

RESULT 27
PLMN RAT STANDARD; PRT; 169 AA.
AC Q01177;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Plasminogen (EC 3.4.21.7) (Fragment).
GN PLG.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1];
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=91250378; PubMed=1645711;
RA Kanalas J.J., Makker S.P.;
RT "Identification of the rat Heymann nephritis autoantigen (GP330) as a
  receptor site for plasminogen."
RL J. Biol. Chem. 266:10823-10829(1991).
CC -!- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as
  a proteolytic factor in a variety of other processes including
  embryonic development, tissue remodeling, tumor invasion, and
  inflammation; in ovulation it weakens the walls of the Graafian
  follicle. It activates the urokinase-type plasminogen activator,
  collagenases and several complement zymogens, such as C1 and C5.
  It cleaves fibrin, fibronectin, thrombospondin, laminin and von
  Willebrand factor.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
  higher selectivity than trypsin. Converts fibrin into soluble
  products.
CC -!- ENZYME REGULATION: Converted into plasmin by plasminogen
  activators, both plasminogen and its activator being bound to
  fibrin. Cannot be activated with streptokinase.
CC -!- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin

```

```

CC immediately after dissociation from the clot.
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
CC -!- SIMILARITY: Contains 5 kringle domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation
  at the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce,
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M62832; AAA41884.1; --
DR PIR; A40522; A40522.
DR HSSP; P00747; 1PMK.
DR MEROPS; S01.233; -.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase S1.
DR Pfam; PF00051; Kringle; 2.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00130; KG; 1.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 2.
DR PROSITE; PS0240; TRYPSIN_DOM; PARTIAL.
DR PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
DR PROSITE; PS00135; TRYPSIN_SER; PARTIAL.
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Repeat.
FT DOMAIN 1 10 KRINGLE 3 (BY SIMILARITY).
FT DOMAIN 34 112 KRINGLE 4 (BY SIMILARITY).
FT DOMAIN 139 >169 KRINGLE 5 (BY SIMILARITY).
FT DISULFID 34 112 BY SIMILARITY.
FT DISULFID 55 95 BY SIMILARITY.
FT DISULFID 83 107 BY SIMILARITY.
FT NON TER 169 169
SQ SEQUENCE 169 AA; 77A54214C49D010C CRC64;

Query Match 29.3%; Score 149.5; DB 1; Length 169;
Best Local Similarity 37.8%; Pred. No. 3.4e-10;
Matches 34; Conservative 12; Mismatches 29; Indels 15; Gaps 6

QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQL---GLGKKNYCRNP 59
Db 34 CYQNGKSYRGTSSTNTCKCOSW-----VSMTPHSHSKTANFPDSSL-ENMYCRNP 87

QY 60 N-RRRPWCYVQGLKPLV--QECMVHDCAD 86
Db 88 NDQRGPWCFT---TDPVVRWEYCNLRKCE 114

RESULT 28
PLMN BOVIN STANDARD; PRT; 812 AA.
AC P06868; Q28162;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Plasminogen precursor (EC 3.4.21.7).
GN PLG.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1];
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Berglund L., Andersen M.D., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen cDNA."

```



Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K.,  
 Masuho Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,  
 Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;  
 "NEDO human cDNA sequencing project";  
 Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 [3]  
 RN SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).  
 RP TISSUE=Brain, and Uterus;  
 RC MEDLINE=22389257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,  
 Diatchenko L., Marasina K., Farmer A.A., Rubin G.M., Hong L.,  
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
 Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 Basak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 Villalón D.K., Muzny D.C., Hales S., Sodergren E.J., Lu X., Gibbs R.A.,  
 Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 Blakeley R.M., Touchman J.W., Green E.D., Dickson M.C.,  
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 Rutterfield V.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
 Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences";  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903 (2002).  
 RL CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf  
 to block Wnt/beta-catenin signaling. Forms a ternary complex with  
 Dkk1 and LRP6 and induces rapid endocytosis and removal of the Wnt  
 receptor LRP6 from the plasma membrane (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC -!- Event=Alternative splicing; Named isoforms=4;  
 CC Name=1;  
 CC IsoId=Q8NCW0-1; Sequence=Displayed;  
 CC Name=2; Synonyms=Kremen2a;  
 CC IsoId=Q8NCW0-2; Sequence=VSP\_050509, VSP\_050510;  
 CC Name=3; Synonyms=Kremen2b;  
 CC IsoId=Q8NCW0-3; Sequence=VSP\_050511, VSP\_050512;  
 CC Name=4; Synonyms=Kremen2c;  
 CC IsoId=Q8NCW0-4; Sequence=VSP\_050513, VSP\_050514;  
 CC -!- SIMILARITY: Contains 1 CUB domain.  
 CC -!- SIMILARITY: Contains 1 kringle domain.  
 CC -!- SIMILARITY: Contains 1 WSC domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See [http://www.ebi.ac.uk/announcements/](http://www.ebi.ac.uk/announcements)  
 or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 DR EMBL; AB086405; BAC00872.1; -;  
 DR EMBL; AB086355; BAC00823.1; -;  
 DR EMBL; AB086356; BAC00824.1; -;  
 DR EMBL; AB086357; BAC00825.1; -;  
 DR EMBL; AK027669; BAB55281.1; -;  
 DR EMBL; AK075033; BAC11365.1; -;  
 DR EMBL; BC003533; AAH03533.1; -;  
 DR EMBL; BC009383; AAH09383.1; -;  
 DR HSSP; P00750; 1PK2.  
 DR Genew; HGNC:18797; KREMEN2.  
 DR InterPro; IPR000859; CUB.  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR002889; WSC.  
 DR Pfam; PF00431; CUB; 1.  
 DR Pfam; PF00051; Kringle; 1.  
 DR Pfam; PF01822; WSC; 1.  
 DR PRINTS; PR00018; KRINGLE.

DR ProDom; PD000395; Kringle; 1.  
 DR SMART; SM00042; CUB; 1.  
 DR SMART; SM00130; KR; 1.  
 DR SMART; SM00321; WSC; 1.  
 DR PROSITE; PS01180; CUB; 1.  
 DR PROSITE; PS00021; KRINGLE\_1; 1.  
 DR PROSITE; PS00070; KRINGLE\_2; 1.  
 DR Wnt signaling pathway; Glycoprotein; Kringle; Signal; Transmembrane;  
 KW Alternative splicing.  
 FT SIGNAL 1 25  
 FT CHAIN 26 462  
 FT KREMEN PROTEIN 2.  
 FT DOMAIN 26 364  
 FT EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 365 387  
 FT POTENTIAL.  
 FT DOMAIN 388 462  
 FT CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 35 119  
 FT KRINGLE.  
 FT DOMAIN 121 215  
 FT WSC.  
 FT DOMAIN 219 326  
 FT CUB.  
 FT CARBOHYD 49 49  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 222 222  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 244 244  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 351 351  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARSPLIC 394 424  
 FT SCLAPGKGFALGASGRFRESWAVYQPPR -> CGALGQ  
 FT GLRADRWGAGAGNARKELGGS (in isoform  
 2).  
 FT /FTId=VSP\_050509.  
 FT Missing (in isoform 2).  
 FT VARSPLIC 425 462  
 FT /FTId=VSP\_050510.  
 FT VARSPLIC 367 420  
 FT ARVFTVTAVSULLLLILLRPIRRRSCLLAPGKPPALG  
 FT ASRGRPRSWAWMY -> GAVCWLRKGRPRWGLPGAPGEAG  
 FT LCGTNSPEGWPCAPPPTPRLLVLPRTGL (in  
 isoform 3).  
 FT /FTId=VSP\_050511.  
 FT Missing (in isoform 3).  
 FT VARSPLIC 421 462  
 FT /FTId=VSP\_050512.  
 FT Missing (in isoform 4).  
 FT ARVFTVTAVSULLLLILLRPIRRRSCLLAP -> GEAG  
 FT ARDGSGRPLAPLITAAVCPQGSRR (in isoform  
 4).  
 FT /FTId=VSP\_050513.  
 FT Missing (in isoform 4).  
 FT VARSPLIC 400 462  
 FT /FTId=VSP\_050514.  
 FT CONFLICT 164 202  
 FT Missing (in Ref. 2; BAC11365).  
 FT CONFLICT 285 285  
 FT A -> D (in Ref. 2; BAC11365).  
 FT SEQUENCE 462 AA; 48849 MW; CE33015917A9AA68 CRC64;  
 SQ  
 Query Match 29.1%; Score 148.5; DB 1; Length 462;  
 Best Local Similarity 42.0%; Pred. No. 1.3e-09;  
 Matches 29; Conservative 9; Mismatches 26; Indels 5; Gaps 2  
 QY 3 CYEGNGHFYRG---KASDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHYCRNPD 59  
 DB 36 CFQVNGADYRGHONRTGPRGAPCLFWDQTO--QHSYSSASDPHGRWGLGHNFCRNP 93  
 QY 60 NRRRPMCYV 68  
 DB 94 GDVQPMCYV 102  
 RESULT 30  
 PLAN ERIEU  
 ID PLMN ERIEU STANDARD; PRT; 810 AA.  
 AC Q294R5;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Plasminogen precursor (EC 3.4.21.7).  
 GN PLG.  
 OS Erinnaceus europaeus (Western European hedgehog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Insectivora; Erinaceidae; Erinaceus.  
 OX NCBI\_TaxID=9365;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;

RA MEDLINE=96025778; PubMed=7592597;  
RA Lawn R.M., Boonmark N.W., Schwartz K., Lindahl G.E., Wade D.P.,  
RA Byrne C.D., Fong K.J., Meer K., Fathiy L.;  
RA "The recurring evolution of lipoprotein(a). Insights from cloning of  
RA hedgehog apolipoprotein(a).";  
RL J. Biol. Chem. 270:24004-24009 (1995).  
RP REVISIONS.  
RA Lawn R.M.;  
RA Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as  
CC a proteolytic factor in a variety of other processes including  
CC embryonic development, tissue remodeling, tumor invasion, and  
CC inflammation; in ovulation it weakens the walls of the Graafian  
CC follicle. It activates the urokinase-type plasminogen activator,  
CC collagenases and several complement zymogens, such as C1 and C5.  
CC It cleaves fibrin, fibronectin, thrombospondin, laminin and von  
CC Willebrand factor.  
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;  
CC higher selectivity than trypsin. Converts fibrin into soluble  
CC products.  
CC -!- ENZYME REGULATION: Converted into plasmin by plasminogen  
CC activators, both plasminogen and its activator being bound to  
CC fibrin. Cannot be activated with streptokinase.  
CC -!- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin  
CC immediately after dissociation from the clot.  
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.  
CC -!- SIMILARITY: Contains 5 kringle domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U33171; AAC48717.1; -  
DR PIR; I46260; I46260.  
DR HSP; P00747; 1PMK.  
DR MEROPS; S01.233; -  
DR InterPro; IPR009003; Cys Ser trypsin.  
DR InterPro; IPR000001; Kringle\_  
DR InterPro; IPR003014; PAN.  
DR InterPro; IPR003609; Pan\_app.  
DR InterPro; IPR01254; Peptidase S1.  
DR InterPro; IPR01314; Peptidase S1A.  
DR InterPro; IPR003966; Peptidase\_S1A\_pr.  
DR Pfam; PF00051; Kringle; 5.  
DR Pfam; PF00024; PAN; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR PRINTS; PR01505; PROTHROMBIN.  
DR ProDom; PD000395; Kringle; 5.  
DR SMART; SM00130; KR; 5.  
DR SMART; SM00473; PAN AP; 1.  
DR SMART; SM00020; Tryp\_Ser; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 5.  
DR PROSITE; PS00070; KRINGLE\_2; 5.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;  
KW Tissue remodeling; blood coagulation; Kringle; Zymogen; Repeat;  
KW Signal.  
FT CHAIN 1 19 BY SIMILARITY.  
FT CHAIN 20 810 PLASMINOGEN.  
FT CHAIN 21 582 PLASMIN HEAVY CHAIN A (BY SIMILARITY).  
FT CHAIN 583 810 PLASMIN LIGHT CHAIN B (BY SIMILARITY).  
FT CHAIN 583 810 SERINE PROTEASE.  
FT DOMAIN 103 181 KRINGLE 1.  
FT DOMAIN 185 262 KRINGLE 2.

FT DOMAIN 275 352 KRINGLE 3.  
FT DOMAIN 379 456 KRINGLE 4.  
FT DOMAIN 482 561 KRINGLE 5.  
FT ACT\_SITE 622 622 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 665 665 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 760 760 CHARGE RELAY SYSTEM.  
FT CARBOHYD 339 339 N-LINKED (GLCNAC... ) (POTENTIAL).  
SQ SEQUENCE 810 AA; 30902 MW; 8E75780946017A16 CRC64;  
  
Query Match 29.1%; Score 148.5; DB 1; Length 810;  
Best Local Similarity 41.5%; Pred. No. 2.3e-09;  
Matches 27; Conservative 7; Mismatches 28; Indels 3; Gaps 1  
  
QY 3 CYEGNGHFYRGKASTDTMGRCLPNNSATVLOQTYHAHRSALQLGLGKHNYCRPDDNR 62  
Db 379 CYQGNQTYRGSSITITGKCPWTSMRPHRSKTPENYPADLTW---NYCRAPDGDK 435  
QY 63 RWCY 67  
Db 436 GPWCY 440  
  
RESULT 31  
PLMN\_PIG  
ID PLMN\_PIG STANDARD; PRT; 790 AA.  
AC P06867;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Plasminogen (EC 3.4.21.7).  
GN PLG.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
[1]  
RP SEQUENCE OF 1-560.  
RA Schaller J., Marti T., Roesselet S.J., Kaempfer U., Rickli E.E.;  
RT "Amino acid sequence of the heavy chain of porcine plasmin. Comparison  
RT of the carbohydrate attachment sites with the human and bovine  
RT species";  
RL Fibrinolysis 1:91-102 (1987).  
RN [2]  
RP SEQUENCE OF 450-790.  
RX MEDLINE=85203907; PubMed=3846533;  
RA Marti T., Schaller J., Rickli E.E.;  
RT "Determination of the complete amino-acid sequence of porcine  
RT miniplasminogen";  
RL Eur. J. Biochem. 149:279-285 (1985).  
RN [3]  
RP CARBOHYDRATE-LINKAGE SITES.  
RX MEDLINE=88185329; PubMed=3356193;  
RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,  
RA Gerwig G.J., van Halbeek H., Vliegenthart J.P.;  
RT "The N- and O-linked carbohydrate chains of human, bovine and porcine  
RT plasminogen. Species specificity in relation to sialylation and  
RT fucosylation patterns";  
RL Eur. J. Biochem. 173:57-63 (1988).  
CC -!- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as  
CC a proteolytic factor in a variety of other processes including  
CC embryonic development, tissue remodeling, tumor invasion, and  
CC inflammation; in ovulation it weakens the walls of the Graafian  
CC follicle. It activates the urokinase-type plasminogen activator,  
CC collagenases and several complement zymogens, such as C1 and C5.  
CC It cleaves fibrin, fibronectin, thrombospondin, laminin and von  
CC Willebrand factor.  
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;  
CC higher selectivity than trypsin. Converts fibrin into soluble  
CC products.  
CC -!- ENZYME REGULATION: Converted into plasmin by plasminogen  
CC activators, both plasminogen and its activator being bound to  
CC fibrin. Cannot be activated with streptokinase.  
CC -!- PTM: N-LINKED GLYCAN CONTAINS N-ACETYLACTOSAMINE, SIALIC ACID AND







RA MEDLINE=20164325; PubMed=10700181;  
RA DeChiara T.M., Kimble R.B., Poueymirou W.T., Rojas J., Masiakowski P.,  
RA Valenzuela D.M., Yancopoulos G.D.;  
RT "Ror2, encoding a receptor-like tyrosine kinase, is required for  
RT cartilage and growth plate development.";  
RL Nat. Genet. 24:271-274(2000).  
CC -!- FUNCTION: Tyrosine-protein kinase receptor which may be involved  
CC in the early formation of the chondrocytes. It seems to be  
CC required for cartilage and growth plate development.  
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. ROR  
CC subfamily.  
CC -!- SIMILARITY: Contains 1 frizzled (FZ) domain.  
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
CC -!- SIMILARITY: Contains 1 kringle domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AB010384; BAA75481.1; -  
CC HSP; P00747; SHPG.  
DR MGI:1347521; Ror2.  
DR GO: GO:0001501; P:skeletal development; IMP.  
DR InterPro: IPR000024; F2\_domain.  
DR InterPro: IPR007110; Ig-like.  
DR InterPro: IPR003598; Ig\_c2.  
DR InterPro: IPR000001; Kringle.  
DR InterPro: IPR00719; Prot\_kinase.  
DR InterPro: IPR001245; Tyr\_kinase.  
DR InterPro: IPR008266; Tyr\_kinase\_AS.  
DR Pfam; PF01392; Fz; 1.  
DR Pfam; PF00047; Ig; 1.  
DR Pfam; PF00051; kringle; 1.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00018; KRINGLE.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000395; Kringle; 1.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR SMART; SM00408; IgC2; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00219; TyRK; 1.  
DR PROSITE; PS50038; FZ; 1.  
DR PROSITE; PS50835; IG LIKE; 1.  
DR PROSITE; PS00021; KRINGLE 1; 1.  
DR PROSITE; PS00070; KRINGLE 2; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; FALSE\_NEG.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
KW Transferrase; Tyrosine-protein kinase; ATP-binding; Receptor;  
KW Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation;  
KW Immunoglobulin domain; Developmental protein.  
FT SIGNAL 1 33  
FT CHAIN 34 944  
FT Tيروسين-PROTEIN KINASE TRANSMEMBRANE  
FT RECEPTOR ROR2.  
FT -----  
FT DOMAIN 34 403  
FT TRANSMEM 404 424  
FT DOMAIN 425 944  
FT DOMAIN 425 944  
FT DOMAIN 55 145  
FT DOMAIN 169 303  
FT DOMAIN 316 394  
FT DOMAIN 473 746  
FT DOMAIN 753 782  
FT DOMAIN 784 857  
FT DOMAIN 859 882  
FT NP\_BIND 479 487  
FT BINDING 507 507  
FT ATP (BY SIMILARITY).  
FT -----

FT ACT\_SITE 615 615 BY SIMILARITY.  
FT MOD\_RES 646 646 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT DISULFID 83 135 BY SIMILARITY.  
FT CARBOHYD 70 70 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 188 188 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 318 318 N-LINKED (GLCNAC...) (POTENTIAL).  
SQ SEQUENCE 944 AA; 105050 MW; CD2EEBC710387A56 CRC64;  
Query Match 28.9%; Score 147.5; DB 1; Length 944;  
Best Local Similarity 41.4%; Pred. No. 3.5e-09;  
Matches 36; Conservative 9; Mismatches 11; Gaps 6  
QY 3 CYEGNGHFYRGKASTDTMGRCPLPWSATVLQQTVAHR---SDALQLGLGHNYCRNPD 59  
DB 316 CYNGSGADYRGWASTTKSGHQCPW---ALQHP-HSHRLSTTEPPELG-GGHAYCRNPG 369  
QY 60 NRRR-PWCYVQVGLPLVQECMVHDC 85  
DB 370 GQMEGFWCFTQ-NKNVRVELCDVPPCS 395  
RESULT 34  
KRM2\_MOUSE  
ID KRM2\_MOUSE STANDARD; PRT; 461 AA.  
AC Q8KIS7;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Kremen protein 2 precursor (Kringle-containing protein marking the eye  
DE and the nose) (Dickkopf receptor 2).  
OS KREMEN2.  
GN Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]\_TaxID=10090;  
RP SEQUENCE FROM N.A., AND FUNCTION.  
RC STRAIN=C57BL/6J;  
RX MEDLINE=22045977; PubMed=12050670;  
RA Mao B., Wu W., Davidson G., Marhold J., Li M., Mechler B.M.,  
RA Delius H., Hoppe D., Stanner P., Walter C., Glinka A., Niehrs C.;  
RT "Kremen proteins are Dickkopf receptors that regulate Wnt/beta-catenin  
RT signalling".  
RL Nature 417:664-667(2002).  
CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf  
CC to block Wnt/beta-catenin signaling. Forms a ternary complex with  
CC Dkk1 and LRP6 and induces rapid endocytosis and removal of the Wnt  
CC receptor LRP6 from the plasma membrane.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).  
CC -!- SIMILARITY: Contains 1 CUB domain.  
CC -!- SIMILARITY: Contains 1 kringle domain.  
CC -!- SIMILARITY: Contains 1 WSC domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AJ457192; CAD29805.1; -  
CC MGD; MGI:1920266; Kremen2.  
DR InterPro: IPR000859; CUB.  
DR InterPro: IPR000001; Kringle.  
DR InterPro: IPR002889; WSC.  
DR Pfam; PF00431; CUB; 1.  
DR Pfam; PF00051; kringle; 1.  
DR Pfam; PF01822; WSC; 1.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00042; CUB; 1.  
DR SMART; SM00130; KR; 1.

```
DR SMART: SM00321; WSC: 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR Wnt signaling pathway; Glycoprotein; Kringle; Signal; Transmembrane.
DR SIGNAL 1 24
DR CHAIN 25 461
DR DOMAIN 25 363
DR TRANSMEM 364 386
DR POTENTIAL.
DR CYTOPLASMIC (POTENTIAL).
DR KRINGLE.
DR WSC.
DR DOMAIN 120 214
DR DOMAIN 218 325
DR CARBOHYD 48 48
DR CARBOHYD 221 221
DR CARBOHYD 243 243
DR CARBOHYD 350 350
DR CARBOHYD 461 AA; 49170 MW; 6DS8C4285809DA CRC64;
SQ SEQUENCE 28.7%; Score 146.5; DB 1; Length 461;
Best Local Similarity 42.0%; Pred. No. 2.2e-09;
Matches 29; Conservative 8; Mismatches 27; Indels 5; Gaps 2;
Qy 3 CYENGHGYRKK---ASTDTGRCPLPWSATVLCQTHAHRSDALQLGKKNYCNPD 59
Db 35 CFQVNGADYRGHONYTGPGRGRCFLFWDQQTQ--QHSYSSASDPQGRWGLGAHNFCTRPD 92
Qy 60 NRRRPWCY 68
Db 93 GDVQFWCY 101
Query Match 28.7%; Score 146.5; DB 1; Length 461;
Best Local Similarity 42.0%; Pred. No. 2.2e-09;
Matches 29; Conservative 8; Mismatches 27; Indels 5; Gaps 2;
CC expressed in fetal and adult CNS and in a variety of human
CC cancers, including those originating from CNS or PNS
CC neuroectoderm.
CC -!- DEVELOPMENTAL STAGE: Expressed at high levels during early
CC embryonic development. The expression levels drop strongly around
CC day 16 and there are only very low levels in adult tissues.
CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. ROR
CC subfamily.
CC -!- SIMILARITY: Contains 1 frizzled (FZ) domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboratio
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on it
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercia
CC entities requires a license agreement (See http://www.isb-sib.ch/announce
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M97675; AAC60275.1; -
CC EMBL; U38894; AAC50714.1; -
CC FIR; A45082; A45082.
CC HSSP; P00747; ICEA.
CC Genew; HGNC:10256; ROR1.
CC MIM; 602336; -
CC GO; GO:0005737; C:cytoplasm; TAS.
CC GO; GO:0005887; C:integral to plasma membrane; TAS.
CC GO; GO:0004714; F:transmembrane receptor protein tyrosine kin. ; TAS.
CC GO; GO:0007169; F:transmembrane receptor protein tyrosine kin. ; TAS.
CC InterPro; IPR000024; Fz_domain.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_C2.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR001245; Tyr_kinase_AS.
CC InterPro; IPR008266; Tyr_kinase_AS.
CC Pfam; PF01392; Fz; 1.
CC Pfam; PF00047; Ig; 1.
CC Pfam; PF00069; kringle; 1.
CC Pfam; PF00051; kinase; 1.
CC PRINTS; PR00108; KRINGLE.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000395; Kringle; 1.
CC ProDom; SM000001; Prot_kinase; 1.
CC SMART; SM00408; IgC2; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00219; Tyrc; 1.
CC PROSITE; PS50038; Fz; 1.
CC PROSITE; PS50835; IG_LIKE; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS50070; KRINGLE_2; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC Transferrase; Tyrosine-protein kinase; ATP-binding; Receptor;
CC Transmembrane; Signal; glycoprotein; Kringle; Phosphorylation;
CC Immunoglobulin domain; Alternative splicing.
CC SIGNAL 1 29
CC CHAIN 30 937
CC DOMAIN 30 406
CC TRANSMEM 407 427
CC DOMAIN 428 937
CC DOMAIN 42 147
CC DOMAIN 165 299
CC DOMAIN 312 391
CC DOMAIN 473 746
CC DOMAIN 784 851
CC DOMAIN 853 876
CC NP_BIND 479 487
CC BINDING 506 506
CC expressed in fetal and adult CNS and in a variety of human
CC cancers, including those originating from CNS or PNS
CC neuroectoderm.
CC -!- DEVELOPMENTAL STAGE: Expressed at high levels during early
CC embryonic development. The expression levels drop strongly around
CC day 16 and there are only very low levels in adult tissues.
CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. ROR
CC subfamily.
CC -!- SIMILARITY: Contains 1 frizzled (FZ) domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboratio
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on it
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercia
CC entities requires a license agreement (See http://www.isb-sib.ch/announce
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M97675; AAC60275.1; -
CC EMBL; U38894; AAC50714.1; -
CC FIR; A45082; A45082.
CC HSSP; P00747; ICEA.
CC Genew; HGNC:10256; ROR1.
CC MIM; 602336; -
CC GO; GO:0005737; C:cytoplasm; TAS.
CC GO; GO:0005887; C:integral to plasma membrane; TAS.
CC GO; GO:0004714; F:transmembrane receptor protein tyrosine kin. ; TAS.
CC GO; GO:0007169; F:transmembrane receptor protein tyrosine kin. ; TAS.
CC InterPro; IPR000024; Fz_domain.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_C2.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR001245; Tyr_kinase_AS.
CC InterPro; IPR008266; Tyr_kinase_AS.
CC Pfam; PF01392; Fz; 1.
CC Pfam; PF00047; Ig; 1.
CC Pfam; PF00069; kringle; 1.
CC Pfam; PF00051; kinase; 1.
CC PRINTS; PR00108; KRINGLE.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000395; Kringle; 1.
CC ProDom; SM000001; Prot_kinase; 1.
CC SMART; SM00408; IgC2; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00219; Tyrc; 1.
CC PROSITE; PS50038; Fz; 1.
CC PROSITE; PS50835; IG_LIKE; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS50070; KRINGLE_2; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC Transferrase; Tyrosine-protein kinase; ATP-binding; Receptor;
CC Transmembrane; Signal; glycoprotein; Kringle; Phosphorylation;
CC Immunoglobulin domain; Alternative splicing.
CC SIGNAL 1 29
CC CHAIN 30 937
CC DOMAIN 30 406
CC TRANSMEM 407 427
CC DOMAIN 428 937
CC DOMAIN 42 147
CC DOMAIN 165 299
CC DOMAIN 312 391
CC DOMAIN 473 746
CC DOMAIN 784 851
CC DOMAIN 853 876
CC NP_BIND 479 487
CC BINDING 506 506
```

FT	ACT_SITE	615	615	BY SIMILARITY.
FT	MOD_RES	645	645	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT	DISULFID	79	131	BY SIMILARITY.
FT	CARBOHYD	47	47	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	66	66	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	184	184	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	315	315	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARSPLIC	1	549	Missing (in isoform Short).
FT				/FTID=VSP_005008.
SEQ	SEQUENCE	937 AA;	104312 MW;	0D0694DBF29F4773 CRC64;
	Query Match	28.6%;	Score 146;	DB 1; Length 937;
	Best Local Similarity	44.1%;	Pred. No. 5.2e-09;	
	Matches 30;	Conservative	5;	Mismatches 25; Indels 8; Gaps 3;
Qy	3	CYEGNGHYFGKASTDMGSPCLPWNASVLTQTYHAHRSALQLG-	IGKKNYCRNDP 60	
Db	313	CYNSTGYDYGTVTSVKRGQCQPMNS-----QYPHTHTFTALRPPELNGHSHYCRNFCN 367		
Qy	61	RRR-PWCY 67		
Db	368	QKEAPWCF 375		
RESULT 36				
ROR1_MOUSE				
AD	ROR1_MOUSE	STANDARD;	PRT;	937 AA.
IC	Q9ZL139;			
DT	16-OCT-2001	(Rel. 40, Created)		
DT	16-OCT-2001	(Rel. 40, Last sequence update)		
DT	10-OCT-2003	(Rel. 42, Last annotation update)		
DE	Tyrosine-protein kinase transmembrane receptor ROR1 precursor			
DE	(EC 2.7.1.112) (Neurotrophic tyrosine kinase, receptor-related 1)			
DE	(mROR1)			
GN	ROR1 OR NTRK1.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A. PubMed=10231392;			
RX	MEDLINE=992484267; PubMed=10231392;			
RA	Oishi I., Takeuchi S., Hashimoto R., Nagabukuro A., Ueda T., Liu Z.J.,			
RA	Hatta T., Akira S., Matsuda Y., Yamamura H., Otani H., Minami Y.;			
RT	"Spatio-temporally regulated expression of receptor tyrosine kinases,			
RT	mRor1, mRor2, during mouse development: implications in development			
RT	and function of the nervous system.";			
RL	Genes Cells 4:41-56 (1999)			
CC	!- FUNCTION: Tyrosine-protein kinase receptor whose role is not yet			
CC	clear.			
CC	!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein			
CC	tyrosine phosphate.			
CC	!- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	!- SIMILARITY: Belongs to the Tyr family of protein kinases. ROR			
CC	subfamily			
CC	!- SIMILARITY: Contains 1 frizzled (FZ) domain.			
CC	!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.			
CC	!- SIMILARITY: Contains 1 kringle domain.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement. See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> .			
CC	EMBL; AB010383; BAA75480.1; --			
DR	HSSP; P00747; ICEA.			
DR	MGD; MGI:1347520; Ror1.			
DR	InterPro; IPR000024; Fz_domain.			
DR	InterPro; IPR007110; Ig-like.			
DR	InterPro; IPR003598; Ig C2.			

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. AND VARIANT ASN-472.  
RX MEDLINE=90202879; PubMed=2318948;  
RA Petersen T.E., Marten M.R., Ichinose A., Davie E.W.;  
RT "Characterization of the gene for human plasminogen, a key proenzyme  
in the fibrinolytic system."  
RL J. Biol. Chem. 265:6104-6111(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87162490; PubMed=3030813;  
RA Forsgren M., Raden B., Israelsson M., Larsson K., Heden L.-O.;  
RT "Molecular cloning and characterization of a full-length cDNA clone  
for human plasminogen."  
RL FEBS Lett. 213:254-260(1987).  
RN [3]  
RP SEQUENCE FROM N.A. AND VARIANTS LYS-57; GLN-133; HIS-261; TRP-409;  
RX ASN-472; VAL-494 AND TRP-523.  
RA Rieder M.J., Arnel T.J., Carrington D.P., Ozuna M., Kuldanek S.A.,  
RA Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE OF 20-810, AND VARIANT ASN-472.  
RA Sottrup-Jensen L., Petersen T.E., Magnusson S.;  
RL Submitted (JUL-1977) to the PIR data bank.  
RN [5]  
RP SEQUENCE OF 292-810 FROM N.A.  
RX MEDLINE=85023311; PubMed=6148961;  
RA Malinowski D.P., Sadler J.E., Davie E.W.;  
RT "Characterization of a complementary deoxyribonucleic acid coding for  
human and bovine plasminogen."  
RL Biochemistry 23:4243-4250(1984).  
RN [6]  
RP SEQUENCE OF 20-100.  
RX MEDLINE=75093329; PubMed=122932;  
RA Wiman B., Wallen P.;  
RT "Structural relationship between 'glutamic acid' and 'lysine' forms  
of human plasminogen and their interaction with the NH2-terminal  
activation peptide as studied by affinity chromatography."  
RL Eur. J. Biochem. 50:489-494(1975).  
RN [7]  
RP SEQUENCE OF 95-580; 581-626; 657-700 AND VARIANT ASN-472.  
RA Sottrup-Jensen L., Claessens H., Zaydel M., Petersen T.E., Magnusson S.;  
RL (In) Davidson J.F., Rowan R.M., Samama M.M., Desnoyers P.C. (eds.);  
RT Progress in chemical fibrinolysis and thrombolysis, pp.3:191-209,  
RL Raven Press, New York (1978).  
RN [8]  
RP SEQUENCE OF 483-604.  
RX MEDLINE=76043692; PubMed=126863;  
RA Wiman B., Wallen P.;  
RT "Amino-acid sequence of the cyanogen-bromide fragment from human  
plasminogen that forms the linkage between the plasmin chains."  
RL Eur. J. Biochem. 58:539-547(1975).  
RN [9]  
RP SEQUENCE OF 581-810.  
RX MEDLINE=77225245; PubMed=142009;  
RA Wiman B.;  
RT "Primary structure of the B-chain of human plasmin."  
RL Eur. J. Biochem. 76:129-137(1977).  
RN [10]  
RP ACTIVE SITE.  
RX MEDLINE=73149248; PubMed=4694729;  
RA Robbins K.C., Bernabe P., Arzadon L., Summaria L.;  
RT "The primary structure of human plasminogen. II. The histidine loop  
of human plasmin; light (B) chain active center histidine sequence."  
RL J. Biol. Chem. 248:1631-1633(1973).  
RN [11]  
RP ACTIVE SITE.  
RX MEDLINE=69234739; PubMed=4240117;  
RA Groskopf W.R., Summaria L., Robbins K.C.;  
RT "Studies on the active center of human plasmin. Partial amino acid  
sequence of a peptide containing the active center serine residue."

J. Biol. Chem. 244:3590-3597(1969).  
RN [12]  
RP OMEGA-AMINOCARBOXYLIC ACID-BINDING SITES.  
RX MEDLINE=82213905; PubMed=6919539;  
RA Trexler M., Vali Z., Paddy L.;  
RT "Structure of the omega-aminocarboxylic acid-binding sites of human  
plasminogen. Arginine 70 and aspartic acid 56 are essential for  
binding of ligand by kringle 4."  
RL J. Biol. Chem. 257:7401-7406(1982).  
RN [13]  
RP FIBRIN AND OMEGA-AMINOCARBOXYLIC ACID BINDING SITES.  
RX MEDLINE=85054794; PubMed=6094526;  
RA Vali Z., Paddy L.;  
RT "The fibrin-binding site of human plasminogen. Arginines 32 and 34  
are essential for fibrin affinity of the kringle 1 domain."  
RL J. Biol. Chem. 259:13690-13694(1984).  
RN [14]  
RP PHOSPHORYLATION SITE SER-597.  
RX MEDLINE=97345939; PubMed=9201958;  
RA Wang H., Prorok M., Bretthauer R.K., Castellino F.J.;  
RT "Serine-578 is a major phosphorylation locus in human plasma  
plasminogen."  
RL Biochemistry 36:8100-8106(1997).  
RN [15]  
RP CARBOHYDRATE-LINKAGE SITES.  
RX MEDLINE=88185329; PubMed=3356193;  
RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,  
RA Gerwig G.J., van Halbeek H., Vliegenthart J.F.;  
RT "The N- and O-linked carbohydrate chains of human, bovine and porcine  
plasminogen. Species specificity in relation to sialylation and  
fucosylation patterns."  
RL Eur. J. Biochem. 173:57-63(1988).  
RN [16]  
RP CARBOHYDRATE-LINKAGE SITE SER-268.  
RX MEDLINE=97207306; PubMed=9054441;  
RA Pirie-Shepherd S.R., Stevens R.D., Andon N.L., Enghild J.J.,  
RA Pizzo S.V.;  
RT "Evidence for a novel O-linked sialylated trisaccharide on Ser-248 of  
human plasminogen 2."  
RL J. Biol. Chem. 272:7408-7411(1997).  
RN [17]  
RP CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.  
RX MEDLINE=95042728; PubMed=7525077;  
RA O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,  
RA Moses M., Lane W.S., Cao E.H., Sage E.H., Folkman J.;  
RT "Angiostatin: a novel angiogenesis inhibitor that mediates the  
suppression of metastases by a Lewis lung carcinoma."  
RL Cell 79:315-328(1994).  
RN [18]  
RP CHARACTERIZATION OF ANGIOSTATIN.  
RX MEDLINE=97238710; PubMed=9102221;  
RA Sim B.K., O'Reilly M.S., Liang H., Fortier A.H., He W., Madsen J.W.,  
RA Lapcevic R., Nacy C.A.;  
RT "A recombinant human angiostatin protein inhibits experimental primary  
and metastatic cancer."  
RL Cancer Res. 57:1329-1334(1997).  
RN [19]  
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 374-461.  
RX MEDLINE=92031502; PubMed=1657148;  
RA Mulichak A.M., Tulinsky A., Ravichandran K.G.;  
RT "Crystal and molecular structure of human plasminogen kringle 4  
refined at 1.9-A resolution."  
RL Biochemistry 30:10576-10588(1991).  
RN [20]  
RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 374-461.  
RX MEDLINE=92031503; PubMed=1657149;  
RA Wu T.-P., Padmanabhan K., Tulinsky A., Mulichak A.M.;  
RT "The refined structure of the epsilon-aminocaproic acid complex of  
human plasminogen kringle 4."  
RL Biochemistry 30:10589-10594(1991).  
RN [21]  
RP X-RAY CRYSTALLOGRAPHY (1.67 ANGSTROMS) OF 376-454.  
RA Stec B., Yamano A., Whitlow M., Teeter M.M.;



```
FT DISULFID 60 66 BY SIMILARITY.
FT DISULFID 110 186 BY SIMILARITY.
FT DISULFID 131 169 BY SIMILARITY.
FT DISULFID 157 181 BY SIMILARITY.
FT DISULFID 191 268 BY SIMILARITY.
FT DISULFID 194 324 INTERCHAIN (BY SIMILARITY).
FT DISULFID 212 251 BY SIMILARITY.
FT DISULFID 240 263 BY SIMILARITY.
FT DISULFID 283 361 BY SIMILARITY.
FT DISULFID 304 343 BY SIMILARITY.
FT DISULFID 332 355 BY SIMILARITY.
FT DISULFID 370 448 BY SIMILARITY.
FT DISULFID 391 431 BY SIMILARITY.
FT DISULFID 419 443 BY SIMILARITY.
FT DISULFID 468 588 INTERCHAIN (BY SIMILARITY).
FT DISULFID 507 523 BY SIMILARITY.
FT DISULFID 602 667 BY SIMILARITY.
FT DISULFID 632 646 BY SIMILARITY.
FT DISULFID 657 685 BY SIMILARITY.
FT CARBOHYD 72 72 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 13 13 Y -> C.
FT VARIANT 212 212 /FTIC=VAR_006631.
FT VARIANT 212 212 C -> F.
FT VARIANT 676 676 /FTIC=VAR_006632.
FT VARIANT 676 676 E -> K (in dbSNP:7798).
FT CONFLICT 623 623 L -> F (in REF. 2).
FT SEQUENCE 711 AA; 80379 MW; 596ED21F180290E4 CRC64;

Query Match 27.9%; Score 142.5; DB 1; Length 711;
Best Local Similarity 33.3%; Pred. No. 9, 9e-09;
Matches 30; Conservative 13; Mismatches 34; Indels 13; Gaps 4;

QY 2 TCYENGHFYRGKASTDTMGRCPLPWSATVLQQTYYAHRSDALQLGLG--KHNYCRNP 59
Db 282 SCFRGKGEGYRGATNTTGTAGVPCQRWDA-----QIPHQRFTPEKYACKOLRENFGRNP 336
QY 60 NRRRPWCYVQVGLKPLVQEC---WVHDCAD 86
Db 337 GSEAPWCFT---LRPGMAAFYQIRRCTD 363

RESULT 39
PLMN CANFA STANDARD; PRT; 333 AA.
AC P80069.
DT 01-NOV-1991 (Rel. 20, Last Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Plasminogen (EC 3.4.21.7) (Fragment).
GN PLG.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxId=9615;
RN [1]
RP SEQUENCE.
RC TISSUE=Plasma;
RX MEDLINE=90175323; PubMed=2626424;
RA Schaller J., Straub C., Kaempfer U., Rickli E.B.;
RL "Complete amino acid sequence of canine miniplasminogen.";
RL Protein Seq. Data Anal. 2:445-450(1989).
CC -!- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as
CC a proteolytic factor in a variety of other processes including
CC embryonic development, tissue remodeling, tumor invasion, and
CC inflammation; in ovulation it weakens the walls of the Graafian
CC follicle. It activates the urokinase-type plasminogen activator,
CC collagenases and several complement zymogens, such as C1 and C5.
CC It cleaves fibrin, fibronectin, thrombospondin, laminin and von
CC Willebrand factor.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
```

```
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC -!- ENZYME REGULATION: Converted into plasmin by plasminogen
CC activators, both plasminogen and its activator being bound to
CC fibrin. Activated with urokinase and high concentrations of
CC streptokinase.
CC -!- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin
CC immediately after dissociation from the clot.
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
CC -!- SIMILARITY: Contains at least 1 kringle domain.
DR HSP; P00747; SHPG.
DR MEROPS; S01.233; -.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase S1A.
DR InterPro; IPR003966; Peptidase_S1A_pr.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen.
FT CHAIN 1 1
FT CHAIN <1 103 PLASMIN HEAVY CHAIN A.
FT CHAIN 104 333 PLASMIN LIGHT CHAIN B.
FT DOMAIN 4 83 KRINGLE 5.
FT DISULFID 104 333 SERINE PROTEASE.
FT DISULFID 4 83 BY SIMILARITY.
FT DISULFID 25 66 BY SIMILARITY.
FT DISULFID 54 78 BY SIMILARITY.
FT DISULFID 90 208 INTERCHAIN (BY SIMILARITY).
FT DISULFID 100 108 INTERCHAIN (BY SIMILARITY).
FT DISULFID 130 146 BY SIMILARITY.
FT DISULFID 222 289 BY SIMILARITY.
FT DISULFID 252 268 BY SIMILARITY.
FT DISULFID 279 307 BY SIMILARITY.
FT ACT_SITE 145 145 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 188 188 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 283 283 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT SITE 152 152 STREPTOKINASE-BINDING SITE (PROBABLE).
FT SITE 186 186 STREPTOKINASE-BINDING SITE (PROBABLE).
FT SITE 264 264 STREPTOKINASE-BINDING SITE (PROBABLE).
FT SITE 277 277 SITE OF SUBSTRATE SPECIFICITY
FT SITE 277 277 (BY SIMILARITY).
SQ SEQUENCE 333 AA; 36678 MW; C8C0271B6C6AC8D4 CRC64;

Query Match 27.6%; Score 141; DB 1; Length 333;
Best Local Similarity 41.4%; Pred. No. 6, 7e-09;
Matches 36; Conservative 5; Mismatches 34; Indels 12; Gaps 5

QY 3 CYEGNGHFYRGKASTDTMGRCPLPWSATVLQQTYYAHR-----SDALQLGLKHYCRNP 58
Db 4 CMFGNGKGYRGKASTDTMGRCPLPWSATVLQQTYYAHR-----QEPHRSIFTPETNPQAGLEK-NYCRNP 57
QY 59 D-NRRRPWCYVQVGLKPLVQECMVHDC 84
Db 58 DGVNPGWCYT-MNQRKLFYCDVPC 83

RESULT 40
PLMN MACMU STANDARD; PRT; 810 AA.
ID PLMN MACMU
AC P12545;
```



DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Plasminogen precursor (EC 3.4.21.7).  
GN PLG.  
OS Macaca mulatta (Rhesus macaque).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecinae; Macaca.  
OX NCBI\_TaxID=9544;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=89174660; PubMed=2925643;  
RA Tomlinson J.E., McLean J.W., Lawn R.M.;  
RT "Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of  
RT synthesis";  
RL J. Biol. Chem. 264:5957-5965(1989).  
CC -!- FUNCTION: plasmin dissolves the fibrin of blood clots and acts as  
CC a proteolytic factor in a variety of other processes including  
CC embryonic development, tissue remodeling, tumor invasion, and  
CC inflammation; in ovulation it weakens the walls of the Graafian  
CC follicle. It activates the urokinase-type plasminogen activator,  
CC collagenases and several complement zymogens, such as C1 and C5.  
CC It cleaves fibrin, fibronectin, thrombospondin, laminin and von  
CC Willebrand factor.  
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;  
CC higher selectivity than trypsin. Converts fibrin into soluble  
CC products.  
CC -!- ENZYME REGULATION: Converted into plasmin by plasminogen  
CC activators, both plasminogen and its activator being bound to  
CC fibrin. Activated with catalytic amounts of streptokinase.  
CC -!- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin  
CC immediately after dissociation from the clot.  
CC -!- MISCELLANEOUS: In the presence of the inhibitor, the activation  
CC involves only cleavage after Arg-580, resulting in 2 chains held  
CC together by 2 disulfide bonds. Without the inhibitor, the  
CC activation involves also removal of the activation peptide.  
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.  
CC -!- SIMILARITY: Contains 5 kringle domains.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL: J04697; AAA36901.1; -  
CC FIR; B32869; B30848.  
CC HSSP; P00747; 1PMK.  
CC MEROPS; S01.233; -  
CC InterPro; IPR009003; Cys\_Ser\_trypsin.  
CC InterPro; IPR000001; Kringle.  
CC InterPro; IPR03014; PAN.  
CC InterPro; IPR03609; Pan\_app.  
CC InterPro; IPR001254; Peptidase\_S1.  
CC InterPro; IPR001314; Peptidase\_S1A.  
CC InterPro; IPR003966; Peptidase\_S1A\_pr.  
CC Pfam; PF00051; kringle; 5.  
CC Pfam; PF00024; PAN\_1.  
CC Pfam; PF00089; trypsin; 1.  
CC PRINTS; PR00722; CHYMOTRYPSIN.  
CC PRINTS; PR00018; KRINGLE.  
CC PRINTS; PR01505; PROTHROMBIN.  
CC ProDom; PD000395; Kringle; 5.  
CC SMART; SM00130; KR; 4.  
CC SMART; SM00473; PAN\_AP; 1.  
CC SMART; SM00020; TRY\_SPC; 1.  
CC PROSITE; PS00021; KRINGLE\_1; 5.  
CC PROSITE; PSS0070; KRINGLE\_2; 5.  
CC PROSITE; PSS0240; TRYPsin\_DOM; 1.  
CC PROSITE; PS00134; TRYPsin\_HIS; 1.

DR PROSITE; PS00135; TRYPsin\_SER; 1.  
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;  
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;  
KW Signal.  
FT SIGNAL 1 19  
FT CHAIN 20 810 PLASMINOGEN.  
FT CHAIN 20 580 PLASMIN HEAVY CHAIN A.  
FT PEPTIDE 20 96 ACTIVATION PEPTIDE.  
FT CHAIN 97 580 PLASMIN SHORT FORM OF CHAIN A.  
FT CHAIN 581 810 PLASMIN LIGHT CHAIN B.  
FT DOMAIN 103 181 KRINGLE 1.  
FT DOMAIN 184 262 KRINGLE 2.  
FT DOMAIN 275 352 KRINGLE 3.  
FT DOMAIN 377 454 KRINGLE 4.  
FT DOMAIN 481 560 KRINGLE 5.  
FT DOMAIN 581 810 SERINE PROTEASE.  
FT ACT\_SITE 622 622 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 665 665 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 760 760 CHARGE RELAY SYSTEM.  
FT BINDING 136 136 OMEGA-AMINOCARBOXYLIC ACIDS.  
FT BINDING 158 158 OMEGA-AMINOCARBOXYLIC ACIDS.  
FT BINDING 172 172 OMEGA-AMINOCARBOXYLIC ACIDS.  
FT BINDING 432 432 OMEGA-AMINOCARBOXYLIC ACIDS.  
FT BINDING 445 445 OMEGA-AMINOCARBOXYLIC ACIDS.  
FT BINDING 134 134 FIBRIN.  
FT BINDING 136 136 FIBRIN.  
FT DISULFID 49 73 BY SIMILARITY.  
FT DISULFID 53 61 BY SIMILARITY.  
FT DISULFID 103 181 BY SIMILARITY.  
FT DISULFID 124 164 BY SIMILARITY.  
FT DISULFID 152 176 BY SIMILARITY.  
FT DISULFID 185 262 BY SIMILARITY.  
FT DISULFID 188 316 BY SIMILARITY.  
FT DISULFID 206 245 BY SIMILARITY.  
FT DISULFID 234 257 BY SIMILARITY.  
FT DISULFID 275 352 BY SIMILARITY.  
FT DISULFID 296 335 BY SIMILARITY.  
FT DISULFID 324 347 BY SIMILARITY.  
FT DISULFID 377 454 BY SIMILARITY.  
FT DISULFID 398 437 BY SIMILARITY.  
FT DISULFID 426 449 BY SIMILARITY.  
FT DISULFID 481 560 BY SIMILARITY.  
FT DISULFID 502 543 BY SIMILARITY.  
FT DISULFID 531 555 BY SIMILARITY.  
FT DISULFID 567 685 BY SIMILARITY.  
FT DISULFID 577 583 BY SIMILARITY.  
FT DISULFID 607 623 BY SIMILARITY.  
FT DISULFID 699 766 BY SIMILARITY.  
FT DISULFID 729 745 BY SIMILARITY.  
FT DISULFID 756 784 BY SIMILARITY.  
FT CARBOHYD 365 365 O-LINKED (GALNAc...) (BY SIMILARITY).  
SQ SEQUENCE 810 AA; 90255 MW; A75E1C51A1A0F24A CRC64;  
Query Match 27.6%; Score 141; DB 1; Length 810;  
Best Local Similarity 37.5%; Pred. No. 1.7e-08;  
Matches 33; Conservative 10; Mismatches 31; Indels 14; Gaps 5  
Qy 3 CYEGNGHFYRGKASTDTMGRPCLPNNSATVLOQTYHAHR---SDALQLGLGKHYCRNPD 59  
Db 377 CYHGDGQSYRGTSSTTTGKKCSWS---MTPWHKTPENFPNAGL-TWNYCRNPD 430  
Qy 60 NRRRPWCYVOGLKPLV--QECMVHCA 85  
Db 431 ADKGPWCFT---TDPVSRWEYCNLKCS 455  
RESULT 41  
FORI\_DROME  
ID\_RORI\_DROME STANDARD; PRT; 685 AA.  
AC Q24488;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Tyrosine-protein kinase transmembrane receptor Roz precursor  
 DE NC 2.7.1.112 (Gkor).  
 GN ROR OR CG4926.  
 GN  
 OS Drosophila melanogaster (Fruit fly).  
 OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OS Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 ON NCBI\_TaxID=7227;  
 RX [1].  
 RP  
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
 RC STRAIN=Canton-S; TISSUE=Larval brain;  
 RC MEDLINE=93348222; PubMed=8394009;  
 RA Wilson C., Goeberdhan D.C.I., Steller H.;  
 RT "Dros. a potential neurotrophic receptor gene, encodes a Drosophila  
 RT homolog of the vertebrate Ror family of Trk-related receptor tyrosine  
 RT kinases.";   
 RL Proc. Natl. Acad. Sci. U.S.A. 90:7109-7113(1993).  
 RP [2].  
 RP SEQUENCE FROM N.A.  
 RP STRAIN=Berkely;  
 RC MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,  
 RA Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Dowes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fodor C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Markulov G., Malshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert X., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue S.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195(2000).  
 RP [3].  
 RP SEQUENCE OF 545-597 FROM N.A.  
 RX MEDLINE=98401146; PubMed=9731193;  
 RA Oates A.C., Wollberg P., Achen M.G., Wilks A.F.;  
 RT "Sampling the genomic pool of protein tyrosine kinase genes using the  
 RT polymerase chain reaction with genomic DNA";  
 RL Biochem. Biophys. Res. Commun. 249:660-667(1998).  
 CC -1- FUNCTION: Tyrosine-protein kinase receptor that functions during  
 CC early stages of neuronal development.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).  
 CC -1- TISSUE SPECIFICITY: Expressed in neurons of the developing nervous  
 CC system.

-!- SIMILARITY: Belongs to the Tyr family of protein kinases. ROR  
subfamily.  
-!- SIMILARITY: Contains 1 frizzled (FZ) domain.  
-!- SIMILARITY: Contains 1 kringle domain.

This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

-----  
EMBL; L20297; AAA28860.1; --  
EMBL; AE003628; AAE52885.1; --  
EMBL; AJ002908; CAA05743.1; --  
PIR; A48289; A48289.  
HSSP; P11362; 1FKG.  
FlyBase; FBgn010407; Ror.  
GO; GO:0016021; C:integral to membrane; NAS.  
GO; GO:0004713; F:protein-tyrosine kinase activity; NAS.  
GO; GO:0007417; P:central nervous system development; IEA.  
GO; GO:0006468; P:protein amino acid phosphorylation; NAS.  
InterPro; IPR000024; Fz\_domain.  
InterPro; IPR000001; Kringle.  
InterPro; IPR000719; Prot\_kinase.  
InterPro; IPR02011; Receptor\_tyr\_kinsII.  
InterPro; IPR001245; Tyr\_pkinase.  
InterPro; IPR008266; Tyr\_pkinase\_AS.  
Pfam; PF00051; Kringle\_1.  
PFAM; PF00069; pkinase; 1.  
PRINTS; PR00018; KRINGLE.  
PRINTS; PR00109; TYRKINASE.  
ProDom; PD000395; Kringle\_1.  
ProDom; PD000001; Prot\_kinase; 1.  
SMART; SMO0130; KR; 1.  
SMART; SMO0219; TyrcK; 1.  
PROSITE; PS00038; FZ; 1.  
PROSITE; PS00021; KRINGLE\_1; 1.  
PROSITE; PS00070; KRINGLE\_2; 1.  
PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
PROSITE; PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
Transferrase; Kinase; Tyrosine-protein kinase; ATP-binding; Receptor;  
Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation;  
Developmental protein.

FT SIGNAL 1 24  
FT CHAIN 25 685

Tyrosine-Protein Kinase Transmembrane  
Receptor ROR.  
EXTRACELLULAR (POTENTIAL).  
POTENTIAL.  
CYTOPLASMIC (POTENTIAL).  
FZ.  
KRINGLE.  
PROTEIN KINASE.  
ATP (BY SIMILARITY).  
ATP (BY SIMILARITY).  
BY SIMILARITY.  
PHOSPHORYLATION (AUTO-) (BY  
SIMILARITY).  
PHOSPHORYLATION (AUTO-) (BY  
SIMILARITY)  
PHOSPHORYLATION (AUTO-) (BY  
SIMILARITY).  
N-LINKED (GLCNAC.. ) (POTENTIAL);  
N-LINKED (GLCNAC.. ) (POTENTIAL);  
N-LINKED (GLCNAC.. ) (POTENTIAL);  
N-LINKED (GLCNAC.. ) (POTENTIAL);  
N-LINKED (GLCNAC.. ) (POTENTIAL);  
N-LINKED (GLCNAC.. ) (POTENTIAL);  
SEQUENCE 685 AA; 78142 MW; 526162D27D5FD7C7 CRC64;

Query Match 27.3%; Score 139; DB 1; Length 685;

Best Local Similarity 34.5%; Pred. No. 2.4e-08;  
Matches 30; Conservative 19; Mismatches 28; Indels 10; Gaps 5;  
Qy 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHYCRNP 60  
Db 235 ENCWEDGTYRGVANVSASGKPLRW--SWLMKEI-----SDPPEL-IGQ-NYCRNPGS 285  
Qy 61 -RRSPWCYVQVGLPLVQECMVHDCAD 86  
Db 286 VENGSPWCFVDSRSRIELCDIPKCAD 312  
RESULT 42  
ID\_HGF\_MOUSE STANDARD; PRT; 728 AA.  
AC Q08048; Q61662; Q64007;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Hepatocyte growth factor precursor (Scatter factor) (SF)  
DE (Hepatopoietin A).  
GN HGF  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT), AND SEQUENCE OF 496-504.  
RC TISSUE=Mammary fibroblast;  
EX MEDLINE=94183257; PubMed=8133822;  
RA Sasaki M., Nishio M., Sasaki T., Enami J.;  
RT "Identification of mouse mammary fibroblast-derived mammary growth  
factor as hepatocyte growth factor.";  
RL Biochem. Biophys. Res. Commun. 199:772-779(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
EX MEDLINE=94363381; PubMed=8081873;  
RA Lee C.C., Kozak C.A., Yamada K.M.;  
RT "Structure, genetic mapping, and expression of the mouse Hgf/scatter  
factor gene.";  
RL Cell Adhes. Commun. 1:101-111(1993).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
EX MEDLINE=94060105; PubMed=8241272;  
RA Liu Y., Michalopoulos G.K., Zarnegar R.;  
RT "Molecular cloning and characterization of cDNA encoding mouse  
hepatocyte growth factor.";  
RL Biochim. Biophys. Acta 1216:299-303(1993).  
CC -!- FUNCTION: HGF is a potent mitogen for mature parenchymal  
hepatocyte cells, seems to be an hepatotrophic factor, and acts as  
growth factor for a broad spectrum of tissues and cell types. It  
has no detectable protease activity.  
CC -!- SUBUNIT: Dimer of an alpha chain and a beta chain linked by a  
disulfide bond.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=Long;  
CC IsoId=Q08048-1; Sequence=Displayed;  
CC Name=Short;  
CC IsoId=Q08048-2; Sequence=VSP\_005408;  
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.  
CC -!- SIMILARITY: Contains 4 kringle domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC -----

DR EMBL; D10212; BAA01064.1; -;  
DR EMBL; D10213; BAA01065.1; -;  
DR EMBL; S71816; AAB31855.1; -;  
DR EMBL; X72307; CAA51054.1; ALT\_INIT.  
DR PIR; JC2117; A60185.  
DR HSSP; P14210; 1BHT.  
DR MEROPS; S01.982; -;  
DR MGD; MGI:96079; Hgf.  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR003014; PAN.  
DR InterPro; IPR003609; Pan\_app.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00051; kringle; 4.  
DR Pfam; PF00024; PAN; 1.  
DR Pfam; PF00089; L\_Typsin; 1.  
DR PRINTS; PR00742; CHYMOTRYPSIN.  
DR PRINTS; PR000018; KRINGLE.  
DR ProDom; PD000395; Kringle; 4.  
DR SMART; SM00130; KR; 4.  
DR SMART; SM00473; PAN AP; 1.  
DR SMART; SM00020; TRYD\_SPC; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 4.  
DR PROSITE; PS00070; KRINGLE\_2; 4.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
KW Growth factor; Kringle; Glycoprotein; Serine protease homolog; Repeat;  
KW Signal; Alternative splicing; Pyrrolidone carboxylic acid.  
FT SIGNAL 1 32 BY SIMILARITY.  
FT CHAIN 33 495 HEPATOCYTE GROWTH FACTOR ALPHA CHAIN.  
FT CHAIN 496 728 HEPATOCYTE GROWTH FACTOR BETA CHAIN.  
FT MOD\_RES 33 33 PYRROLIDONE CARBOXYLIC ACID  
(BY SIMILARITY).  
FT DOMAIN 33 128 PAP.  
FT DOMAIN 129 207 KRINGLE 1.  
FT DOMAIN 212 289 KRINGLE 2.  
FT DOMAIN 306 384 KRINGLE 3.  
FT DOMAIN 392 470 KRINGLE 4.  
FT DOMAIN 496 728 SERINE PROTEASE-LIKE.  
FT DISULFID 71 97 BY SIMILARITY.  
FT DISULFID 75 85 BY SIMILARITY.  
FT CARBOHYD 488 607 INTERCHAIN (BY SIMILARITY).  
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 569 569 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 656 656 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VARSPIC 163 167 Missing (in isoform Short).  
FT CONFLICT 344 344 /FTIGVSP\_005408.  
FT CONFLICT 479 479 N -> K (IN REF. 2).  
FT CONFLICT 564 564 V -> L (IN REF. 2).  
FT CONFLICT 564 564 R -> H (IN REF. 3).  
SQ SEQUENCE 728 AA; 82944 MW; A0381FC497534328 CRC64;  
Query Match 27.0%; Score 137.5; DB 1; Length 728;  
Best Local Similarity 32.0%; Pred. No. 3.8e-08;  
Matches 31; Conservative 12; Mismatches 35; Indels 19; Gaps 4  
Qy 3 CYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLG----KHNYCRNP 58  
Db 306 CIQGGEGYRGNTIWNIGIPQCRWDS-----QYPKHDIITPENFKCKDLRENYCRNP 358  
Qy 59 DNRRRPWCY-----VQVGLKPLVQECMV---HDCADG 87  
Db 359 DGAESPWCFTTDPNIRVGYCSQIPKCDVSSGQDCYRG 395  
RESULT 43  
HGF\_RAT STANDARD; PRT; 728 AA.  
ID\_HGF\_RAT  
AC P17945;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hepatocyte growth factor precursor (Scatter factor) (SF)  
DE (Hepatopoietin A).  
GN HGF.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
[1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC STRAIN=WiStar; TISSUE=Liver;  
RX MEDLINE=90222197; PubMed=2139229;  
RA Shimizu S., Nakamura T.,  
RA "Deduced primary structure of rat hepatocyte growth factor and  
RT expression of the mRNA in rat tissues".  
RL Proc. Natl. Acad. Sci. U.S.A. 87:3200-3204(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WiStar; TISSUE=Liver;  
RX MEDLINE=91031482; PubMed=2146117;  
RA Okajima A., Miyazawa K., Kitamura N.;  
RT "Primary structure of rat hepatocyte growth factor and induction of  
RT its mRNA during liver regeneration following hepatic injury.";  
RL Eur. J. Biochem. 193:375-381(1990).  
CC -!- FUNCTION: HGF is a potent mitogen for mature parenchymal  
CC hepatocyte cells, seems to be an hepatotrophic factor, and acts as  
CC growth factor for a broad spectrum of tissues and cell types. It  
CC has no detectable protease activity  
CC -!- SUBUNIT: Dimer of an alpha chain and a beta chain linked by a  
CC disulfide bond.  
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.  
CC -!- SIMILARITY: Contains 4 kringle domains.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
DR EMBL; D90102; BAA14133.1; -;  
DR EMBL; X54400; CAA38266.1; -;  
DR PIR; A35644; A35644.  
DR HSSP; F14210; 1BHT.  
DR MEROPS; S01.978; -;  
DR InterPro; IPR009003; Cys Ser trypsin.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR003014; PAN.  
DR InterPro; IPR003609; Pan.app.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00051; kringle; 4.  
DR Pfam; PF00024; PAN; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR PRODOM; PD000395; Kringle; 4.  
DR SMART; SM00130; KR; 4.  
DR SMART; SM00473; PAN AP; 1.  
DR SMART; SM00020; Tryp Spc; 1.  
DR PROSITE; PS00021; KRINGLE 1; 4.  
DR PROSITE; PS00070; KRINGLE 2; 4.  
DR PROSITE; PS50240; TRYPSIN\_DOM; 1.  
KW Growth factor; Kringle; Glycoprotein; Serine protease homolog; Repeat;  
KW Signal; Pyrrolidone carboxylic acid.  
FT SIGNAL 1 32 BY SIMILARITY.  
FT CHAIN 33 495 HEPATOCYTE GROWTH FACTOR ALPHA CHAIN.  
FT CHAIN 496 728 HEPATOCYTE GROWTH FACTOR BETA CHAIN.  
FT MOD\_RES 33 33 PYRROLIDONE CARBOXYLIC ACID  
FT (BY SIMILARITY).  
FT DOMAIN 33 128 PAP.  
FT DOMAIN 129 207 KRINGLE 1.

FT DOMAIN 212 289 KRINGLE 2.  
FT DOMAIN 306 384 KRINGLE 3.  
FT DOMAIN 392 470 KRINGLE 4.  
FT DOMAIN 496 728 SERINE PROTEASE-LIKE.  
FT DISULFID 71 97 BY SIMILARITY.  
FT DISULFID 75 85 BY SIMILARITY.  
FT DISULFID 498 607 INTERCHAIN (BY SIMILARITY).  
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 569 569 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 656 656 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 728 AA; 82905 NW; 3E0BF1F96ADCEDF6 CRC64;  
Query Match 26.6%; Score 135.5; DB 1; Length 728;  
Best Local Similarity 32.0%; Pred. No. 6.5e-08;  
Matches 31; Conservative 11; Mismatches 36; Indels 19; Gaps 4  
QY 3 CYEKGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSALQLGLG----KHNYCRNP 58  
DB 306 CIKGQGEYRGTTNTIWNIGIPQCRWDS-----QYPHKHDITPENFKDLRENYCRNP 358  
QY 59 DNRRRPWCY-----VOVGLKPLVQECMV---HDCADG 87  
DB 359 DGAESFWCFTDPNIRVGYCSQIPKCDVSSGQDCYRG 395  
RESULT 44  
HGF\_HUMAN STANDARD; PRT; 728 AA.  
ID HGF\_HUMAN AC F14210; Q9BYL9; Q9UDU6;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hepatocyte growth factor precursor (Scatter factor) (SF)  
DE (Hepatopoietin A).  
GN Homo sapiens (Human).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91340155; PubMed=1831432;  
RA Seki T., Hagiya M., Shimonishi M., Nakamura T., Shimizu S.;  
RT "Organization of the human hepatocyte growth factor-encoding gene.";  
RL Gene 102:213-219(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RX MEDLINE=89392017; PubMed=2528952;  
RA Miyazawa K., Tsubouchi H., Naka D., Takahashi K., Okigaki M.,  
RA Arakaki N., Nakayama H., Hirono S., Sakiyama O., Takahashi K.,  
RA Gohda E., Daikuhara Y., Kitamura N.;  
RT "Molecular cloning and sequence analysis of cDNA for human hepatocyte  
RT growth factor.";  
RL Biochem. Biophys. Res. Commun. 163:967-973(1989).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Leukocyte;  
RX MEDLINE=91025062; PubMed=2145836;  
RA Seki T., Ihara I., Sugimura A., Shimonishi M., Nishizawa T.,  
RA Asami O., Hagiya M., Nakamura T., Shimizu S.;  
RT "Isolation and expression of cDNA for different forms of hepatocyte  
RT growth factor from human leukocyte.";  
RL Biochem. Biophys. Res. Commun. 172:321-327(1990).  
RN [4]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 55-73 AND 495-520.  
RC TISSUE=Liver;  
RX MEDLINE=90066676; PubMed=2531289;  
RA Nakamura T., Nishizawa T., Hagiya M., Seki T., Shimonishi M.,  
RA Sugimura A., Tashiro K., Shimizu S.;  
RT "Molecular cloning and expression of human hepatocyte growth factor.";  
RL Nature 342:440-443(1999).

RN RP SEQUENCE FROM N.A.  
RX TISSUE=Embryonic fibroblast;  
RX MEDLINE=91334393; PubMed=1831266;  
RA Rieder M.J., Livingston R.J., Daniels M.R., Chung M.-W.,  
RA Rieder M.J., Arakaki N., Hartmann G., Vandekerckhove J., Weingart S.,  
RA Rieder H., Fonatsch C., Teubouchi H., Hishida T., Daikuhara Y.,  
RA Birchemier W.,  
RT Evidence for the identity of human scatter factor and human  
RT hepatocyte growth factor.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:7001-7005 (1991).  
[6]  
RN RP SEQUENCE FROM N.A.  
RA Rieder M.J., Livingston R.J., Daniels M.R., Chung M.-W.,  
RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,  
RA Schackwitz W.S., Sherwood J.K., Wittrak L.A., Nickerson D.A.;  
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
[7]  
RN RP SEQUENCE FROM N.A.  
RA Courtney L., Elliot G., Angell S.;  
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
[8]  
RN RP SEQUENCE OF 249-695 FROM N.A.  
RX MEDLINE=91369928; PubMed=1832556;  
RA Miyazawa K., Kitamura A., Kitamura N.;  
RT "Structural organization and the transcription initiation site of the  
RT human hepatocyte growth factor gene.";  
RL Biochemistry 30:9170-9176 (1991).  
[9]  
RN RP SIGNAL SEQUENCE CLEAVAGE SITE  
RX MEDLINE=91207365; PubMed=1826837;  
RA Yoshiyama Y., Arakaki N., Naka D., Takahashi K., Hirono S., Kondo J.,  
RA Nakayama H., Gohda E., Kitamura N., Tsubouchi H., Ishii T.,  
RA Hishida T., Daikuhara Y.;  
RT Identification of the N-terminal residue of the heavy chain of both  
RT native and recombinant human hepatocyte growth factor.";  
RL Biochem. Biophys. Res. Commun. 175:660-667 (1991).  
[10]  
RN RP CARBOHYDRATE-LINKAGE SITE THR-476  
RX MEDLINE=93129192; PubMed=1482348;  
RA Shimizu N., Hara H., Sogabe T., Sakai H., Ihara I., Inoue H.,  
RA Nakamura T., Shimizu S.;  
RT "Hepatocyte growth factor is linked by O-glycosylated oligosaccharide  
RT on the alpha chain.";  
RL Biochem. Biophys. Res. Commun. 189:1329-1335 (1992).  
[11]  
RN RP MUTAGENESIS  
RX MEDLINE=92331602; PubMed=1321034;  
RA Lokker N.A., Mark M.R., Luis E.A., Bennett G.L., Robbins K.A.,  
RA Baker J.B., Godowski P.J.;  
RT "Structure-function analysis of hepatocyte growth factor:  
RT identification of variants that lack mitogenic activity yet retain  
RT high affinity receptor binding.";  
RL EMBO J. 11:2503-2510 (1992).  
[12]  
RN RP STRUCTURE BY NMR OF 31-127  
RX MEDLINE=98154323; PubMed=9493272;  
RA Zhou H., Mazzulla M.J., Kaufman J.D., Stahl S.J., Wingfield P.T.,  
RA Rubin J.S., Sottaro D.P., Byrd R.A.;  
RT "The solution structure of the N-terminal domain of hepatocyte growth  
RT factor reveals a potential heparin-binding site.";  
RL Structure 6:1109-1116 (1998).  
[13]  
RN RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 35-210  
RX MEDLINE=99036858; PubMed=9817840;  
RA Ulsch M., Lokker N.A., Godowski P.J., de Vos A.M.;  
RT "Crystal structure of the N-terminal fragment of human hepatocyte growth  
RT factor at 2.0-A resolution.";  
RL Structure 6:1383-1393 (1998).  
CC CC -!- FUNCTION: HGF is a potent mitogen for mature parenchymal  
CC hepatocyte cells, seems to be an hepatotrophic factor, and acts  
CC as growth factor for a broad spectrum of tissues and cell types.  
CC It has no detectable protease activity.  
CC -!- SUBUNIT: Dimer of an alpha chain and a beta chain linked by a

CC CC disulfide bond.  
CC CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.  
CC CC -!- SIMILARITY: Contains 4 kringle domains.  
CC CC  
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC CC  
CC EMBL; D90334; BAAL4348.1; -  
DR EMBL; D90318; BAAL4348.1; JOINED.  
DR EMBL; D90319; BAAL4348.1; JOINED.  
DR EMBL; D90320; BAAL4348.1; JOINED.  
DR EMBL; D90322; BAAL4348.1; JOINED.  
DR EMBL; D90323; BAAL4348.1; JOINED.  
DR EMBL; D90324; BAAL4348.1; JOINED.  
DR EMBL; D90325; BAAL4348.1; JOINED.  
DR EMBL; D90326; BAAL4348.1; JOINED.  
DR EMBL; D90327; BAAL4348.1; JOINED.  
DR EMBL; D90328; BAAL4348.1; JOINED.  
DR EMBL; D90329; BAAL4348.1; JOINED.  
DR EMBL; D90330; BAAL4348.1; JOINED.  
DR EMBL; D90331; BAAL4348.1; JOINED.  
DR EMBL; D90332; BAAL4348.1; JOINED.  
DR EMBL; D90333; BAAL4348.1; JOINED.  
DR EMBL; D90334; BAAL4348.1; JOINED.  
DR EMBL; M29145; AAAS2650.1; -  
DR EMBL; M60718; AAAS2648.1; -  
DR EMBL; X16323; CAA34387.1; -  
DR EMBL; M73239; AAA64239.1; -  
DR EMBL; M73240; AAA64297.1; -  
DR EMBL; AX246560; AAO61091.1; -  
DR EMBL; AC004960; AAC71655.1; -  
DR EMBL; M75983; AAG53460.1; -  
DR EMBL; M75972; AAG53460.1; JOINED.  
DR EMBL; M75973; AAG53460.1; JOINED.  
DR EMBL; M75974; AAG53460.1; JOINED.  
DR EMBL; M75975; AAG53460.1; JOINED.  
DR EMBL; M75976; AAG53460.1; JOINED.  
DR EMBL; M75977; AAG53460.1; JOINED.  
DR EMBL; M75978; AAG53460.1; JOINED.  
DR EMBL; M75979; AAG53460.1; JOINED.  
DR EMBL; M75980; AAG53460.1; JOINED.  
DR EMBL; M75981; AAG53460.1; JOINED.  
DR EMBL; M75982; AAG53460.1; JOINED.  
DR PIR; JH0579; JH0579.  
DR PDB; 2HGF; 24-JUN-98.  
DR PDB; 1BHT; 18-NOV-98.  
DR PDB; INKI; 29-DEC-99.  
DR MEROPS; S01.976; -.  
DR GlycoSuiteDB; P14210; -.  
DR Genew; HGNC:4893; HGF.  
DR MIM; 142409; -.  
DR GO; GO:0008083; F: growth factor activity; NAS.  
DR GO; GO:0007067; P: mitosis; NAS.  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR003014; PAN.  
DR InterPro; IPR003609; Pan app.  
DR InterPro; IPR001254; Peptidase S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00051; kringle; 4.  
DR Pfam; PF00024; PAN; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000395; Kringle; 4.  
DR SMART; SM00130; KR; 4.  
DR SMART; SM00473; PAN AP; 1.  
DR SMART; SM00020; Tryp\_Spc; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 4.

DR PROSITE; PS50070; KRINGLE 2; 4.  
 DR PROSITE; PS50240; TRYPSIN\_DOM; 1.  
 KW Growth factor; Kringle; Glycoprotein; Serine protease homolog; Repeat;  
 KW Signal; 3D-structure; Polymorphism; Pyrrolidone carboxylic acid.  
 FT SIGNAL 1 31  
 FT CHAIN 32 494 HEPATOCYTE GROWTH FACTOR ALPHA CHAIN.  
 FT CHAIN 495 728 HEPATOCYTE GROWTH FACTOR BETA CHAIN.  
 FT MOD RES 32 32 PYRROLIDONE CARBOXYLIC ACID.  
 FT DOMAIN 32 127 PAP.  
 Query Match 26.2%; Score 133.5; DB 1; Length 728;  
 Best Local Similarity 34.7%; Pred. No. 1.1e-07;  
 Matches 33; Conservative 10; Mismatches 37; Indels 15; Gaps 5;  
 QY 3 CYEGNGHYRGKASTDTMGRPCLPWNSATVLQQTYYAHF--RSDALQLGLGKHNYCRNPDN 60  
 Db 305 CIQGGEGYRGVNTIWMGIFQQRWDS-----QYPHEHDMTPENFKCKDLRENYCRNPDG 359  
 QY 61 RRRPKY-----VQGLXPLVQEC-WVH--DCADG 87  
 Db 360 SESPWFTTDNIRVGYCSQIPNCDMSHGQDCYRG 394  
 RESULT 45  
 HGFL MOUSE  
 ID HGFL MOUSE STANDARD; PRT; 716 AA.  
 AC P24928;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Hepatocyte growth factor-like protein precursor (Macrophage  
 DE stimulatory protein) (MSP).  
 GN MST1 OR HGFL.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CC NCBI\_TaxID=10090;  
 CC [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c; TISSUE=Liver;  
 RX MEDLINE=92002017; PubMed=1832957;  
 RA Fritzen Degen S.J., Stuart L.A., Han S., Jamison C.S.;  
 RT "Characterization of the mouse cDNA and gene coding for a hepatocyte  
 RT growth factor-like protein: expression during development.";  
 RL Biochemistry 30:9781-9791(1991).  
 CC -!- FUNCTION: Probably has no proteolytic activity, since crucial AA  
 CC characteristic of serine proteases catalytic sites are not  
 CC conserved.  
 CC -!- TISSUE SPECIFICITY: Liver. Lower levels in lung, placenta and  
 CC adrenal.  
 CC -!- DEVELOPMENTAL STAGE: Is expressed at low levels during gestation.  
 CC Just before birth the level increases dramatically and remains  
 CC stable afterwards.  
 CC -!- PTM: MAY BE CLEAVED AFTER AA 488, TO YIELD A TWO-CHAIN MOLECULE  
 CC HELD TOGETHER BY DISULFIDE BONDS, OR CLEAVED INTO TWO SEPARATE  
 CC POLYPEPTIDES.  
 CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.  
 CC -!- SIMILARITY: Contains 4 kringle domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M74180; AAA50166.1; -;  
 CC EMBL; M74181; AAA50167.1; -;  
 CC FIR; A40332; A40332.  
 CC HSRP; P00747; 1KEN.  
 CC MEROPS; S01.975; -;  
 CC MGD; MGI:96080; Mst1.

DR GO; GO:0007566; P.embryo implantation; IC.  
 DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR003014; PAN.  
 DR InterPro; IPR003609; Pan app.  
 DR InterPro; IPR001254; Peptidase S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR InterPro; IPR003966; Peptidase\_S1A\_pr.  
 DR Pfam; PF00051; Kringle; 4.  
 DR Pfam; PF00024; PAN; 1.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PRINTS; PR0018; KRINGLE.  
 DR PRINTS; PR01505; PROTHROMBIN.  
 DR ProDom; PD000395; Kringle; 4.  
 DR SMART; SM00130; KR; 4.  
 DR SMART; SM00473; PAN AP; 1.  
 DR SMART; SM00020; TRYPSIN; 1.  
 DR PROSITE; PS00021; KRINGLE 1; 4.  
 DR PROSITE; PS0070; KRINGLE 2; 4.  
 DR PROSITE; PS50240; TRYPSIN\_DOM; 1.  
 KW Kringle; Glycoprotein; Serine protease homolog; Repeat; Signal.  
 FT SIGNAL 1 18 POTENTIAL.  
 FT CHAIN 19 716 HEPATOCYTE GROWTH FACTOR-LIKE PROTEIN.  
 FT DOMAIN 19 109 PAP.  
 FT DOMAIN 110 186 KRINGLE 1.  
 FT DOMAIN 191 268 KRINGLE 2.  
 FT DOMAIN 292 370 KRINGLE 3.  
 FT DOMAIN 379 457 KRINGLE 4.  
 FT DOMAIN 489 716 SERINE PROTEASE-LIKE.  
 FT DISULFID 56 78 BY SIMILARITY.  
 FT DISULFID 60 66 BY SIMILARITY.  
 FT DISULFID 110 186 BY SIMILARITY.  
 FT DISULFID 131 169 BY SIMILARITY.  
 FT DISULFID 157 181 BY SIMILARITY.  
 FT DISULFID 191 268 BY SIMILARITY.  
 FT DISULFID 194 333 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 212 251 BY SIMILARITY.  
 FT DISULFID 240 283 BY SIMILARITY.  
 FT DISULFID 292 370 BY SIMILARITY.  
 FT DISULFID 313 352 BY SIMILARITY.  
 FT DISULFID 341 364 BY SIMILARITY.  
 FT DISULFID 379 457 BY SIMILARITY.  
 FT DISULFID 400 440 BY SIMILARITY.  
 FT DISULFID 428 452 BY SIMILARITY.  
 FT DISULFID 477 593 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 512 528 BY SIMILARITY.  
 FT DISULFID 607 672 BY SIMILARITY.  
 FT DISULFID 637 651 BY SIMILARITY.  
 FT DISULFID 662 690 BY SIMILARITY.  
 FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 173 173 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 305 305 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 620 620 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CONFLICT 19 19 P -> Q (IN GENOMIC SEQUENCE).  
 SQ SEQUENCE 716 AA; 80588 MW; BBCE0EF85213ACC CRC64;  
 Query Match 25.9%; Score 132; DB 1; Length 716;  
 Best Local Similarity 33.8%; Pred. No. 1.6e-07;  
 Matches 27; Conservative 11; Mismatches 22; Indels 20; Gaps 4.  
 QY 3 CYEGNGHYRGKASTDTMGRPCLPWNSATVLQQTYYAHF-----SDALQLGLGKHNY 54  
 Db 292 CFRGKGEDYRGTTNTTSAGVPCQRMDA-----QSPHGRFVPEKYACKDL-----RENF 340  
 QY 55 CRNPDNRPRPCVY-QVGLK 73  
 Db 341 CRNPDGSEAPWCFTSRPGLR 360  
 RESULT 46  
 PLMN MOUSE  
 ID PLMN\_MOUSE STANDARD; PRT; 812 AA.

AC P20918; Q8C182; Q91WJ5;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].  
GN PLG.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91184812; PubMed=2081500;  
RA Degen S.J., Bell S.M., Schaefer L.A., Elliott R.W.;  
RT "Characterization of the cDNA coding for mouse plasminogen and  
RT localization of the gene to mouse chromosome 17.";  
RL Genomics 8:49-61(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX STRAIN=129/Sv;  
RA Brathwaite M., Waeltz P., Qian Y., Dudekula D., Schlessinger D.,  
RA Nagaraja R.;  
RT "Genomic sequence analysis in the mouse t-complex region.";  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Liver;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J.J., Heide F.,  
RA Diachenko L., Maruina K.B., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullah S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Heiton E., Kettner M., Madan A.A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Buterfield V.S.N., Krzyzinski M.T., Skalska U., Smalus D.B.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [4]  
RP SEQUENCE OF 1-16 FROM N.A.  
RX STRAIN=129/SvJ; TISSUE=Liver;  
RX PubMed=12149246;  
RA Bannach F.G., Gutierrez A., Fowler B.J., Bugge T.H., Degen J.L.,  
RA Farmer R.J., Miles L.A.;  
RT "Localization of regulatory elements mediating constitutive and  
RT cytokine-stimulated plasminogen gene expression.";  
RL J. Biol. Chem. 277:38579-38588(2002).  
RN [5]  
RP CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.  
RX MEDLINE=95042728; PubMed=7525077;  
RA O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,  
RA Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;  
RT "Angiostatin: a novel angiogenesis inhibitor that mediates the  
RT suppression of metastases by a Lewis lung carcinoma.";  
RL Cell 79:315-328(1994).  
CC -!- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as  
CC a proteolytic factor in a variety of other processes including  
CC embryonic development, tissue remodeling, tumor invasion, and  
CC inflammation; in ovulation it weakens the walls of the Graafian  
CC follicle. It activates the urokinase-type plasminogen activator,  
CC collagenases and several complement zymogens, such as C1 and C5.  
CC It cleaves fibrin, fibronectin, thrombospondin, laminin and von  
CC Willebrand factor.

CC -!- FUNCTION: Angiostatin is an angiogenesis inhibitor that blocks  
CC neovascularization and growth of experimental primary and  
CC metastatic tumors in vivo.  
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;  
CC higher selectivity than trypsin. Converts fibrin into soluble  
CC products.  
CC -!- ENZYME REGULATION: Converted into plasmin by plasminogen  
CC activators, both plasminogen and its activator being bound to  
CC fibrin. Cannot be activated with streptokinase.  
CC -!- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin  
CC immediately after dissociation from the clot.  
CC -!- MISCELLANEOUS: In the presence of the inhibitor, the activation  
CC involves only cleavage after Arg-581, resulting in 2 chains held  
CC together by 2 disulfide bonds. Without the inhibitor, the  
CC activation involves also removal of the activation peptide.  
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.  
CC -!- SIMILARITY: Contains 5 kringle domains.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collabora-  
CC tion between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercia-  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: J04766; AA850168.1; -;  
CC EMBL: AF481053; AM22155.1; -;  
CC EMBL: BC014773; AAH14773.1; -;  
CC EMBL: BC057186; AAH57186.1; -;  
CC EMBL: AY134430; AAH15805.1; -;  
CC F1R; A38514; F1MS.  
CC HSSP; P00747; 1PMK.  
CC MEROPS; S01.233; -;  
CC MGD; MGI:97620; Plg.  
CC GO; GO:0016506; Frapoptosis activator activity; IDA.  
CC GO; GO:0008915; Frapoptosis; IDA.  
CC InterPro; IPR009003; Cys Ser trypsin.  
CC InterPro; IPR000001; Kringle.  
CC InterPro; IPR003014; PAN.  
CC InterPro; IPR003609; Pan app.  
CC InterPro; IPR001254; Peptidase S1.  
CC InterPro; IPR001314; Peptidase\_S1A.  
CC InterPro; IPR003966; Peptidase\_S1A pr.  
CC Pfam; PF00051; Kringle; 5.  
CC Pfam; PF00024; PAN; 1.  
CC Pfam; PF00089; trypsin; 1.  
CC PRINTS; PR00722; CHYMOTRYPSIN.  
CC PRINTS; PR00018; KRINGLE.  
CC PRINTS; PR01505; PROTHROMBIN.  
CC ProDom; PD000395; Kringle; 4.  
CC SMART; SM00130; KR; 5.  
CC SMART; SM00473; PAN AP; 1.  
CC SMART; SM00020; Tryp SPC; 1.  
CC PROSITE; PS00021; KRINGLE\_1; 5.  
CC PROSITE; PS00070; KRINGLE\_2; 5.  
CC PROSITE; PS00240; TRYPSIN\_DOM; 1.  
CC PROSITE; PS00334; TRYPSIN\_HIS; 1.  
CC PROSITE; PS00135; TRYPSIN\_SER; 1.  
CC KW Hydrolyase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;  
CC KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;  
CC KW Signal.  
FT CHAIN 1 19 PLASMINOGEN.  
FT CHAIN 20 812 PLASMIN HEAVY CHAIN A.  
FT PEPTIDE 20 581 ACTIVATION PEPTIDE.  
FT CHAIN 98 97 PLASMIN SHORT FORM OF CHAIN A.  
FT CHAIN 98 7436 ANGIOSTATIN.  
FT CHAIN 582 812 PLASMIN LIGHT CHAIN B.  
FT DOMAIN 103 181 KRINGLE 1.  
FT DOMAIN 184 262 KRINGLE 2.  
FT DOMAIN 275 352 KRINGLE 3.  
FT DOMAIN 377 454 KRINGLE 4.







CC STRAIN=FVB/N; TISSUE=Liver;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,  
RA Altschul S.F., Zeberg B., Busto K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsien F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Locuelli N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs S.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield V.S.N., Krzyzinski M.I., Skalska U., Smallus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [3]  
RP SEQUENCE OF 384-618 FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=92212913; PubMed=1557383;  
RA Banfield D.K., Macgillivray R.T.,  
RT "Partial characterization of vertebrate prothrombin cDNAs:  
RT amplification and sequence analysis of the B chain of thrombin from  
RT nine different species";  
RL Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783 (1992).  
CC -!- FUNCTION: Thrombin, which cleaves bonds after Arg and Lys,  
CC converts fibrinogen to fibrin and activates factors V, VII, VIII,  
CC XIII, and in complex with thrombomodulin, protein C,  
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Gly; activates  
CC fibrinogen to fibrin and releases fibrinopeptide A and B.  
CC -!- PTM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,  
CC RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOmal  
CC ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES  
CC ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY  
CC CHARGED PHOSPHOLIPID SURFACE, WHICH IS ESSENTIAL FOR THE CONVERSION  
CC OF PROTHROMBIN TO THROMBIN.  
CC -!- MISCELLANEOUS: Prothrombin is activated on the surface of a  
CC phospholipid membrane that binds the amino end of prothrombin and  
CC factors Va and Xa in Ca-dependent interactions; factor Xa removes  
CC the activation peptide and cleaves the remaining part into light  
CC and heavy chains. The activation process starts slowly because  
CC factor V itself has to be activated by the initial, small amounts  
CC of thrombin.  
CC -!- MISCELLANEOUS: Thrombin can itself cleave the amino terminal  
CC fragment (fragment 1) of the prothrombin, prior to its activation  
CC by factor Xa.  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
CC -!- SIMILARITY: Contains 2 kringle domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X52308; CAA36548.1; -;  
CC EMBL; BC013662; AAH13662.1; -;  
CC EMBL; M81394; AAA40435.1; -;  
CC PIR; A35827; A35827.  
CC HGSP; P00734; I87X.  
CC MEROPS; S01.217; -;  
CC MGD; MGI:88380; F2.  
CC InterPro; IPR009003; Cys\_Ser\_trypsin.  
CC InterPro; IPR002383; GLA\_Blood.  
CC InterPro; IPR000001; Kringle.

DR InterPro; IPR001354; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR InterPro; IPR003966; Peptidase\_S1A\_pr.  
DR InterPro; IPR000294; VitK\_dep\_GLA.  
DR Pfam; PF00594; Gla; 1.  
DR Pfam; PF00051; Kringle; 2.  
DR Pfam; PF00089; Trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00001; GLABLOOD.  
DR PRINTS; PR00018; KRINGLE.  
DR PRINTS; PR01505; PROTHROMBIN.  
DR ProDom; PD000395; Kringle; 2.  
DR SMART; SM00069; GLA; 1.  
DR SMART; SM00130; KR; 2.  
DR SMART; SM00020; TRYPSIN; 1.  
DR PROSITE; PS00011; GLD CARBOXYLATION; 1.  
DR PROSITE; PS00021; KRINGLE 1; 2.  
DR PROSITE; PS00070; KRINGLE 2; 2.  
DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Blood coagulation; Plasma; Calcium-binding; Glycoprotein; Repeat;  
KW Vitamin K; zymogen; Gamma-carboxyglutamic acid; Acute phase; Liver;  
KW Hydrolyase; Serine protease; Kringle; Signal.  
KW SIGNAL 1 24  
FT PROPEP 25 43 PROTHROMBIN.  
FT CHAIN 44 618  
FT PEPTIDE 44 200 ACTIVATION PEPTIDE (FRAGMENT 1).  
FT PEPTIDE 201 324 ACTIVATION PEPTIDE (FRAGMENT 2).  
FT CHAIN 325 360 THROMBIN LIGHT CHAIN (A).  
FT CHAIN 361 618 THROMBIN HEAVY CHAIN (B).  
FT DOMAIN 109 187 KRINGLE 1.  
FT DOMAIN 215 292 KRINGLE 2.  
FT DOMAIN 361 618 SERINE PROTEASE.  
FT SITE 200 201 CLEAVAGE (BY THROMBIN).  
FT SITE 324 325 CLEAVAGE (BY FACTOR XA).  
FT SITE 360 361 CLEAVAGE (BY FACTOR XA).  
FT ACT\_SITE 403 403 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 459 459 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 565 565 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT MOD\_RES 50 50 GAMMA-CARBOXYGLUTAMIC ACID.  
FT MOD\_RES 51 51 GAMMA-CARBOXYGLUTAMIC ACID.  
FT MOD\_RES 58 58 GAMMA-CARBOXYGLUTAMIC ACID.  
FT MOD\_RES 60 60 GAMMA-CARBOXYGLUTAMIC ACID.  
FT MOD\_RES 63 63 GAMMA-CARBOXYGLUTAMIC ACID.  
FT MOD\_RES 64 64 GAMMA-CARBOXYGLUTAMIC ACID.  
FT MOD\_RES 69 69 GAMMA-CARBOXYGLUTAMIC ACID.  
FT MOD\_RES 70 70 GAMMA-CARBOXYGLUTAMIC ACID.  
FT MOD\_RES 73 73 GAMMA-CARBOXYGLUTAMIC ACID.  
FT MOD\_RES 76 76 GAMMA-CARBOXYGLUTAMIC ACID.  
FT DISULFID 61 66 BY SIMILARITY.  
FT DISULFID 91 104 BY SIMILARITY.  
FT DISULFID 109 187 BY SIMILARITY.  
FT DISULFID 130 170 BY SIMILARITY.  
FT DISULFID 158 182 BY SIMILARITY.  
FT DISULFID 215 293 BY SIMILARITY.  
FT DISULFID 236 276 BY SIMILARITY.  
FT DISULFID 264 288 BY SIMILARITY.  
FT DISULFID 333 479 INTERCHAIN (BY SIMILARITY).  
FT DISULFID 388 404 BY SIMILARITY.  
FT DISULFID 533 547 BY SIMILARITY.  
FT DISULFID 561 591 BY SIMILARITY.  
FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .).  
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .).  
FT CARBOHYD 413 413 N-LINKED (GLCNAC. . .).  
FT CARBOHYD 553 553 N-LINKED (GLCNAC. . .).  
SQ SEQUENCE 618 AA; 70268 MW; B89F719A6FD601E0 CRC64;  
Query Match 24.6%; Score 125.5; DB 1; Length 618;  
Best Local Similarity 31.8%; Pred. No. 7.8e-07;  
Matches 27; Conservative 13; Mismatches 40; Indels 5; Gaps 3  
QY 3 CYEGNGHYRGKASTDTNGRPCLPWN SATVLQTYHAHRSDALQLGLGKNYCENPD-NR 61

Db 215 CLTERGLYQGNLAVTTIGSFCLPWNLPKATLSKYQDFPEVKL---VENFCRNFDDWE 271  
 QY 62 RRPWCYVQVGLKPLVQECWHDCA 86  
 Db 272 EGAWCYV-AGOPGDFEYCNLYCEE 295

RESULT 49  
 ID THRB HUMAN STANDARD; PRT; 622 AA.  
 AC P00734;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Prothrombin precursor (EC 3.4.21.5) (Coagulation factor II).  
 GN F2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88077877; PubMed=2825773;  
 RA Degen S.J.F., Davie E.W.;  
 RT "Nucleotide sequence of the gene for human prothrombin.";  
 RL Biochemistry 26:6165-6177(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND VARIANT MET-165.  
 RA Rieder M.J., Arnel T.Z., Carrington D.P., Chung M.-W., Lee K.L.,  
 RA Ozuna M., Poel C.L., Toth E.J., Yi Q., Nickerson D.A.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 8-622 FROM N.A.  
 RX MEDLINE=83231469; PubMed=6305407;  
 RA Degen S.J.F., McGillivray R.T.A., Davie E.W.;  
 RT "Characterization of the complementary deoxyribonucleic acid and gene  
 coding for human prothrombin.";  
 RL Biochemistry 22:2087-2097(1983).  
 RN [4]  
 RP SEQUENCE OF 44-314.  
 RX MEDLINE=77193964; PubMed=266717;  
 RA Walz D.A., Hewett-Smyth D., Seegers W.H.;  
 RT "Amino acid sequence of human prothrombin fragments 1 and 2.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 74:1969-1972(1977).  
 RN [5]  
 RP SEQUENCE OF 315-622.  
 RX MEDLINE=77207112; PubMed=873923;  
 RA Burkowski R.J., Elion J., Downing M.R., Mann K.G.;  
 RT "Primary structure of human prothrombin 2 and alpha-thrombin.";  
 RL J. Biol. Chem. 252:4942-4957(1977).  
 RN [6]  
 RP PROCESSING.  
 RX MEDLINE=87008532; PubMed=3759958;  
 RA Rabiet M.J., Blashill A., Furie B., Furie B.C.;  
 RT "Prothrombin fragment 1 X 2 X 3, a major product of prothrombin  
 activation in human plasma.";  
 RL J. Biol. Chem. 261:13210-13215(1986).  
 RN [7]  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
 RX MEDLINE=90059942; PubMed=2583108;  
 RA Bode W., Mayr I., Baumann U., Huber R., Stone S.R., Hofsteenge J.;  
 RT "The refined 1.9 Å crystal structure of human alpha-thrombin:  
 interaction with D-Phe-Pro-Arg chloromethylketone and significance of  
 the Tyr-Pro-Pro-Trp insertion segment.";  
 RL EMBO J. 8:3467-3475(1989).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
 RX MEDLINE=90327074; PubMed=2374926;  
 RA Rydel T.J., Ravichandran K.G., Tullinsky A., Bode W., Huber R.,  
 RA Roitsch C., Fenton J.W. II;  
 RT "The structure of a complex of recombinant hirudin and human alpha-  
 thrombin.";

Science 249:277-280(1990).  
 RN [9]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).  
 RX MEDLINE=94350942; PubMed=8071320;  
 RA Rydel T.J., Vin M., Padmanabhan K.P., Blankenship D.T., Cardin A.D.,  
 RA Correa P.E., Fenton J.W. II, Tullinsky A.;  
 RT "Crystallographic structure of human gamma-thrombin.";  
 RL J. Biol. Chem. 269:22000-22006(1994).  
 RN [10]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
 RX MEDLINE=97357286; PubMed=9214615;  
 RA van de Loch A., Bode W., Huber R., le Bonniec B.F., Stone S.R.,  
 RA Esmen C.T., Stubbs M.T.;  
 RT "The thrombin E192Q-BPTI complex reveals gross structural  
 rearrangements: implications for the interaction with antithrombin  
 and thrombomodulin.";  
 RL EMBO J. 16:2977-2984(1997).  
 RN [11]  
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 328-601.  
 RX MEDLINE=99162521; PubMed=10051558;  
 RA Quinto E.R., Caccia S., Rose T., Fueterer K., Waksman G., di Cera E.;  
 RT "Unexpected crucial role of residue 225 in serine proteases.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:1852-1857(1999).  
 RN [12]  
 RP VARIANT DYSPROTHROMBINEMIA CYS-314.  
 RX MEDLINE=87033739; PubMed=3771562;  
 RA Rabiet M.-J., Furie B.C., Furie B.;  
 RT "Molecular defect of prothrombin Barcelona. Substitution of cysteine  
 for arginine at residue 273.";  
 RL J. Biol. Chem. 261:15045-15048(1986).  
 RN [13]  
 RP VARIANT DYSPROTHROMBINEMIA ALA-509.  
 RX MEDLINE=95313001; PubMed=7792730;  
 RA Degen S.J.F., McDowell S.A., Sparks L.M., Scharer I.;  
 RT "Prothrombin Frankfurt: a dysfunctional prothrombin characterized by  
 substitution of Glu-466 by Ala.";  
 RL Thromb. Haemost. 73:203-209(1995).  
 RN [14]  
 RP VARIANTS DYSPROTHROMBINEMIA THR-380 AND HIS-431.  
 RX MEDLINE=93043342; PubMed=1421398;  
 RA Morishita E., Saito M., Kumabashiri I., Asakura H., Matsuda T.,  
 RA Yamaguchi K.;  
 RT "Prothrombin Himi: a compound heterozygote for two dysfunctional  
 prothrombin molecules (Met-337-->Thr and Arg-388-->His)."  
 RL Blood 80:2275-2280(1992).  
 RN [15]  
 RP VARIANT DYSPROTHROMBINEMIA HIS-314.  
 RX MEDLINE=95169898; PubMed=7865694;  
 RA James H.L., Kim D.J., Zheng D.-Q., Girolami A.;  
 RT "Prothrombin Padua I: incomplete activation due to an amino acid  
 substitution at a factor Xa cleavage site.";  
 RL Blood Coagul. Fibrinolysis 5:841-844(1994).  
 RN [16]  
 RP VARIANT DYSPROTHROMBINEMIA CYS-425.  
 RX MEDLINE=89207504; PubMed=3242619;  
 RA Henriksen R.A., Mann K.G.;  
 RT "Identification of the primary structural defect in the dysfibrinogen  
 thrombin Quick I: substitution of cysteine for arginine-382.";  
 RL Biochemistry 27:9160-9165(1988).  
 RN [17]  
 RP VARIANT DYSPROTHROMBINEMIA VAL-601.  
 RX MEDLINE=89247398; PubMed=2719946;  
 RA Henriksen R.A., Mann K.G.;  
 RT "Substitution of valine for glycine-558 in the congenital dysfibrinogen  
 thrombin Quick II alters primary substrate specificity.";  
 RL Biochemistry 28:2078-2082(1989).  
 RN [18]  
 RP VARIANT DYSPROTHROMBINEMIA ALA-509.  
 RX MEDLINE=92378975; PubMed=1354985;  
 RA Miyata T., Aruga R., Uneyama H., Bezeaud A., Guillain M.-C.,  
 RA Iwanaga S.;  
 RT "Prothrombin Salakta: substitution of glutamic acid-466 by alanine  
 reduces the fibrinogen clotting activity and the esterase activity.";

Biochemistry 31:7457-7462(1992).  
[19]  
RN VARIANT DYSPROTHROMBINEMIA TRP-461.  
RP MEDLINE=87185407; PubMed=3567158;  
RX Miyata T., Morita T., Inomoto T., Kawachi S., Shirakami A.,  
RA Iwanaga S.;  
RT "Prothrombin Tokushima, a replacement of arginine-418 by tryptophan  
RT that impairs the fibrinogen clotting activity of derived thrombin  
RT Tokushima.";  
RN Biochemistry 26:1117-1122(1987).  
[20]  
RN VARIANT DYSPROTHROMBINEMIA TRP-461.  
RP MEDLINE=87101511; PubMed=3801671;  
RX Inomoto T., Shirakami A., Kawachi S.,  
RA Miyoshi K., Morita T., Iwanaga S.;  
RT "Prothrombin Tokushima: characterization of dysfunctional thrombin  
RT derived from a variant of human prothrombin.";  
RN Blood 69:565-569(1987).  
[21]  
RN VARIANT DYSPROTHROMBINEMIA TRP-461.  
RP MEDLINE=92256895; PubMed=1349838;  
RX Iwahana H., Yoshimoto K., Shigekiyo T., Shirakami A., Saito S.,  
RA Itakura M.;  
RT "Detection of a single base substitution of the gene for prothrombin  
RT Tokushima. The application of PCR-SSCP for the genetic and molecular  
RT analysis of dysprothrombinemia.";  
RN Int. J. Hematol. 55:93-100(1992).  
[22]  
RN VARIANT DYSPROTHROMBINEMIA LYS-200.  
RP MEDLINE=83204687; PubMed=6405779;  
RX Board P.G., Shaw D.C.;  
RA "Determination of the amino acid substitution in human prothrombin  
RT type 3 (157 Glu leads to Lys) and the localization of a third  
RT thrombin cleavage site.";  
RN Br. J. Haematol. 54:245-254(1983).  
[23]  
RN VARIANTS MET-165 AND THR-386.  
RX MEDLINE=993118093; PubMed=10391209;  
RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
RA Shaw N., Lane C.R., Lim E.P., Kalyanaram N., Nimesh J., Ziaugra L.,  
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,  
RA Lander E.S.;  
RN Nat. Genet. 23:373-373(1999).  
CC -!- FUNCTION: Thrombin, which cleaves bonds after Arg and Lys,  
CC converts fibrinogen to fibrin and activates factors V, VII, VIII,  
CC XIII, and, in complex with thrombomodulin, protein C.  
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Gly; activates  
CC fibrinogen to fibrin and releases fibrinopeptide A and B.  
CC -!- SUBCELLULAR LOCATION: Extracellular.  
CC -!- TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.  
CC -!- PTM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,  
CC RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOmal  
CC ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES  
CC ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY  
CC CHARGED PHOSPHOLIPID SURFACE, WHICH IS ESSENTIAL FOR THE CONVERSION

Query Match 24.3%; Score 124; DB 1; Length 622;  
Best Local Similarity 38.8%; Pred. No. 1.2e-06;  
Matches 26; Conservative 8; Mismatches 29; Indels 4; Gaps 2;  
QY 3 CYESGNHFVYRGASTDTMGRCLPWNATVLTQOYTHAHRSDALQGLGKHNYCRNPD-NR 61  
DB 213 CVPDRGQQYQGRGLAVTTHTGLPLCLAWASQAQKALSKHQDFNSAVQL---VENFCRNPDPGE 269

QY 62 RRPWCYV 68  
DB 270 EGVWCYV 276  
RESULT 50  
PLMN SHEEP  
ID PLMN SHEEP STANDARD; PRT; 343 AA.  
AC P81256;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Plasminogen (EC 3.4.21.7) (Fragment).  
GN PLG.  
OS Ovis aries (Sheep).  
OC Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;  
OC Bovidae; Caprinae; Ovis.  
OC NCBI\_TaxID=9940;  
OX [1]  
RN SEQUENCE.  
RX MEDLINE=93149995; PubMed=1492092;  
RA Schaller J., Straub C., Kampfer U., Rickli E.E.;  
RT "Complete amino acid sequence of ovine miniplasminogen.";  
RN Protein Seq. Data Anal. 5:21-25(1992).  
CC -!- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as  
CC a proteolytic factor in a variety of other processes including  
CC embryonic development, tissue remodeling, tumor invasion, and  
CC inflammation; in ovulation it weakens the walls of the Graafian  
CC follicle. It activates the urokinase-type plasminogen activator,  
CC collagenases and several complement zymogens, such as C1 and C5.  
CC It cleaves fibrin, fibronectin, thrombospondin, laminin and von  
CC Willebrand factor.  
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;  
CC higher selectivity than trypsin. Converts fibrin into soluble  
CC products.  
CC -!- ENZYME REGULATION: Converted into plasmin by plasminogen  
CC activators, both plasminogen and its activator being bound to  
CC fibrin. Cannot be activated with streptokinase.  
CC -!- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin  
CC immediately after dissociation from the clot.  
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.  
CC -!- SIMILARITY: Contains at least 2 kringle domains.  
DR PIR; B61545; B61545.  
DR HSP; P00747; SHPG.  
DR MEROPS; S01.233;  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR InterPro; IPR003966; Peptidase\_S1A\_pr.  
DR Pfam; PF00051; kringle; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR PRINTS; PR01505; PROTHROMBIN.  
DR PRODOM; PD000395; Kringle; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS50070; KRINGLE\_2; 1.  
DR PROSITE; PS02440; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;  
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat.  
FT NON\_TER 1 1  
FT DOMAIN <1 140 HEAVY CHAIN A.  
FT DOMAIN 141 >343 LIGHT CHAIN A.  
FT DOMAIN <1 17 KRINGLE 4.  
FT DOMAIN 114 120 KRINGLE 5.  
FT DOMAIN 114 341 SERINE PROTEASE.  
FT ACT\_SITE 181 CHARGE RELAY SYSTEM.

FT ACT\_SITE 224 224 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 319 319 CHARGE RELAY SYSTEM.  
FT NON\_TER 343 343  
SQ SEQUENCE 343 AA; 37662 MW; 8DF6EBA92D596E0 CRC64;  
  
Query Match 23.1%; Score 118; DB 1; Length 343;  
Best Local Similarity 36.8%; Pred. No. 3.1e-06;  
Matches 32; Conservative 6; Mismatches 37; Indels 12; Gaps 5;  
  
Qy 3 CYEGNGHFYRGKASTDTMGRPCLPWSATVLQOTYHAH-----RSDALQLGLGKHNYCRNP 58  
Db 15 CMLGIGKGYRGKATTVAGVPCQEWAA-----QEPHRHGIFTETNPRAGLEX-NYCRNP 69  
  
Qy 59 D-NRRRPWCYVQVGLKPLVQECMVHDC 84  
Db 69 DGDVNGPWCYT-TNPRKLFDCIPQC 94

Search completed: May 25, 2004, 14:53:48  
Job time : 5.39351 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2004, 14:44:05 ; Search time 14.5767 Seconds

(without alignments)  
1904.795 Million cell updates/sec

Title: US-09-880-503-1  
Perfect score: 510  
Sequence: 1 KTCVGNHGHEFVGKASTDTM.....QVGLKPLVQECMVHDCADGK 88

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

- Database :
- 1: SP\_ARCHAEA:\*
  - 2: SP\_BACTERIA:\*
  - 3: SP\_FUNGI:\*
  - 4: SP\_HUMAN:\*
  - 5: SP\_INVERTEBRATE:\*
  - 6: SP\_MAMMAL:\*
  - 7: SP\_MHC:\*
  - 8: SP\_ORGANELLE:\*
  - 9: SP\_PHAGE:\*
  - 10: SP\_PLANT:\*
  - 11: SP\_RODENT:\*
  - 12: SP\_VIRUS:\*
  - 13: SP\_VERTEBRATE:\*
  - 14: SP\_UNCLASSIFIED:\*
  - 15: SP\_VIRUS:\*
  - 16: SP\_BACTERIAP:\*
  - 17: SP\_ARCHAEP:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	495	97.1	154	4 Q96SE8	Q96se8 homo sapien
2	415	81.4	433	6 QMIL0	Qmil0 oryctolagus
3	415	81.4	433	6 QMIL0	Qmil0 oryctolagus
4	396	77.6	157	6 Q8VAB	Q8vab bos taurus
5	387	75.9	231	11 Q8C6L2	Q8c6l2 mus musculus
6	313	61.4	214	6 Q8XT70	Q8xt70 oryctolagus
7	235	45.1	564	6 Q8MKB1	Q8mb1 oryctolagus
8	226	44.3	291	4 Q7Z7N2	Q7z7n2 homo sapien
9	226	44.3	516	4 Q8BU99	Q8bu99 homo sapien
10	226	44.3	562	4 Q8EYK8	Q8eyk8 homo sapien
11	218.5	42.8	90	4 Q8NG20	Q8ng20 homo sapien
12	215	42.2	395	4 Q8BZW1	Q8bzw1 homo sapien
13	210	41.2	562	6 Q8SQ23	Q8sq23 sus scrofa
14	209.5	41.1	517	11 Q8K0D2	Q8k0d2 mus musculus
15	202	39.6	540	13 Q800X7	Q800x7 meleagris g
16	193.5	37.9	560	4 Q14520	Q14520 homo sapien

17	191	37.5	202	13	Q90675	Q90675 gallus gall
18	188	36.9	653	11	Q8VCS4	Q8vcs4 mus musculus
19	186.5	36.6	597	11	Q35727	Q35727 mus musculus
20	186.5	36.6	609	11	Q80YC5	Q80yc5 mus musculus
21	177.5	34.8	616	6	Q97507	Q97507 sus scrofa
22	172.5	33.8	421	13	Q8AXX3	Q8axx3 xenopus lae
23	170.5	33.4	615	4	Q8IZZ5	Q8izz5 homo sapien
24	163.5	32.1	947	13	Q8AXY6	Q8axy6 gallus gall
25	158	31.0	454	6	Q45506	Q45506 papio hamad
26	156	30.6	113	4	Q9UIR5	Q9uir5 homo sapien
27	155.5	30.5	263	13	Q7SXB3	Q7sxb3 brachydanio
28	154	30.2	806	6	Q18783	Q18783 macropus eu
29	153	30.0	113	4	Q9UIR7	Q9uir7 homo sapien
30	151.5	29.7	359	6	Q8WMR1	Q8wmr1 canis famli
31	149.5	29.3	812	11	Q9R0K3	Q9r0k3 rattus norv
32	149	29.2	716	13	Q31691	Q31691 xenopus lae
33	148.5	29.1	801	11	Q8K0Q8	Q8k0q8 mus musculus
34	148.5	29.1	944	11	Q8C3W2	Q8c3w2 mus musculus
35	148.5	29.1	944	11	Q8ESP6	Q8esp6 mus musculus
36	148	29.0	105	4	Q9UIR8	Q9uir8 homo sapien
37	147	28.8	113	4	Q9UIR6	Q9uir6 homo sapien
38	147	28.8	648	4	Q9H1V4	Q9h1v4 homo sapien
39	146	28.6	393	4	Q9BRB6	Q9brb6 homo sapien
40	146	28.6	567	4	Q13208	Q13208 homo sapien
41	146	28.6	937	11	Q8BNP9	Q8bnp9 mus musculus
42	146	28.6	937	11	Q8BG10	Q8bg10 mus musculus
43	144	28.2	810	4	Q15146	Q15146 homo sapien
44	143.5	28.1	709	13	Q90ZN6	Q90zn6 brachydanio
45	143	28.0	902	5	Q17576	Q17576 caenorhabdi
46	143	28.0	928	5	Q9BLV1	Q9blv1 caenorhabdi
47	140.5	27.5	759	11	Q7TP84	Q7tp84 rattus norv
48	140	27.5	132	4	Q16609	Q16609 homo sapien
49	140	27.5	930	13	Q8AV69	Q8av69 xenopus lae
50	138.5	27.2	704	13	Q90865	Q90865 gallus gall
51	138.5	27.2	709	13	Q7ZTN9	Q7ztn9 xenopus lae
52	138.5	27.2	717	13	P70006	P70006 xenopus lae
53	138	27.1	1145	5	Q9BKL8	Q9bkl8 aplysia cal
54	137.5	27.0	513	13	Q8FUF8	Q8fuf8 crocodylus
55	137.5	27.0	728	11	Q8C9G5	Q8c9g5 mus musculus
56	135.5	26.6	553	13	Q7Z322	Q7z322 brachydanio
57	134.5	26.4	730	6	Q867B7	Q867b7 canis famli
58	132.5	26.0	215	11	Q81123	Q81123 mus musculus
59	132.5	26.0	215	13	Q42341	Q42341 gallus gall
60	132.5	26.0	264	11	Q81122	Q81122 mus musculus
61	132.5	26.0	264	11	Q7TMJ8	Q7tmj8 mus musculus
62	132.5	26.0	726	13	Q90978	Q90978 gallus gall
63	132.5	26.0	2869	6	Q28398	Q28398 erinaceus e
64	132	25.9	716	11	Q91XG8	Q91xg8 mus musculus
65	131	25.7	95	4	Q8N696	Q8n696 homo sapien
66	131	25.7	145	6	Q28911	Q28911 macaca fasc
67	131	25.7	234	4	Q85YW2	Q85yw2 homo sapien
68	131	25.7	263	4	Q00318	Q00318 homo sapien
69	131	25.7	263	4	Q8NCJ9	Q8ncj9 homo sapien
70	131	25.7	263	4	Q96FE7	Q96fe7 homo sapien
71	131	25.7	716	11	P70521	P70521 rattus norv
72	130	25.5	334	6	Q46507	Q46507 papio hamad
73	129	25.3	429	13	Q8AVB0	Q8avb0 brachydanio
74	129	25.3	594	5	P91823	P91823 caenorhabdi
75	127.5	25.0	710	13	Q91402	Q91402 xenopus. he

ALIGNMENTS

RESULT 1  
Q96SE8 PRELIMINARY; PRT; 154 AA.  
ID Q96SE8;  
AC Q96SE8;  
DT 01-DEC-2001 (T-EMBLrel. 19, Created)  
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)  
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)  
DS Urokinase-type plasminogen activator amino-terminal fragment  
GN AIF.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Fu J., Bai X., Ruan C.;  
 RT "Cloning and expression of the amino-terminal fragment of human  
 RT urokinase-type plasminogen activator.";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Bai X., Fu J., Wang W., Xi X., Ruan C.;  
 RT "Overexpression of the amino-terminal fragment of human urokinase-type  
 RT plasminogen activator in breast cancer cells results in decreased  
 RT tumor invasion, growth and angiogenesis.";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  
 DR EMBL; AY029537; AAK38734.1; -.  
 DR GO; GO:0016301; F:kinase activity; IEA.  
 DR InterPro; IPR006209; EGF-like.  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR000001; Kringle.  
 DR Pfam; PF00051; kringle; 1.  
 DR ProDom; PD000395; Kringle; 1.  
 DR SMART; SM00181; EGF; 1.  
 DR SMART; SM00130; KR; 1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS00021; KRINGLE\_1; 1.  
 DR PROSITE; PS00070; KRINGLE\_2; 1.  
 KW Glycoprotein; Kinase; Kringle.  
 SQ SEQUENCE 154 AA; 17305 MW; A3CCF2PCFF505572 CRC64;  
 Query Match 97.1%; Score 495; DB 4; Length 154;  
 Best Local Similarity 98.9%; Pred. No. 5.8e-50;  
 Matches 86; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQQTTHAHRSDALQLGLGKHNCRNPDN 60  
 DB 68 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQQTTHAHRSDALQLGLGKHNCRNPDN 127  
 QY 61 RRPWCYVQVGLKPLVQECWVHDCADG 87  
 DB 128 RRPWCYVQVGLKPLVQECWVHDCADG 154  
 RESULT 2  
 ID Q8M1L0 PRELIMINARY; PRT; 433 AA.  
 AC Q8M1L0;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Urokinase-type plasminogen activator.  
 GN PLAU.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=2215945; PubMed=12149463;  
 RA Falkenberg M., Tom C., DeYoung M.B., Wen S., Linnemann R.,  
 RA Dichek D.A.;  
 RT "Increased expression of urokinase during atherosclerotic lesion  
 RT development causes arterial constriction and lumen loss, and  
 RT accelerates lesion growth.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:10665-10670(2002).  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  
 DR EMBL; AY122285; AAM83187.1; -.  
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.

DR GO; GO:0016301; F:kinase activity; IEA.  
 DR GO; GO:0008233; F:peptidase activity; IEA.  
 DR GO; GO:0004295; F:trypsin activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR009003; Cys Ser trypsin.  
 DR InterPro; IPR006209; EGF-like.  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR InterPro; IPR008293; Pept\_S1A\_UPA.  
 DR Pfam; PF00051; kringle; 1.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PRINTS; PR00018; KRINGLE.  
 DR ProDom; PD000395; Kringle; 1.  
 DR SMART; SM00130; KR; 1.  
 DR SMART; SMC0020; Tryp\_Spc; 1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS00021; KRINGLE\_1; 1.  
 DR PROSITE; PS00070; KRINGLE\_2; 1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 DR PIRSF; PIRSF001144; Ufk\_plasm\_act; 1.  
 KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.  
 SQ SEQUENCE 433 AA; 48375 MW; 65E64F36415549B0 CRC64;  
 Query Match 81.4%; Score 415; DB 6; Length 433;  
 Best Local Similarity 79.5%; Pred. No. 3.9e-40;  
 Matches 70; Conservative 7; Mismatches 11; Indels 0; Gaps 0  
 QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQQTTHAHRSDALQLGLGKHNCRNPDN 60  
 DB 70 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQQTTHAHRSDALQLGLGKHNCRNPDN 129  
 QY 61 RRPWCYVQVGLKPLVQECWVHDCADG 88  
 DB 130 RRPWCYVQVGLKPLVQECWVHDCADG 157  
 RESULT 3  
 ID Q8MHY7 PRELIMINARY; PRT; 433 AA.  
 AC Q8MHY7;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Urokinase-type plasminogen activator.  
 GN UROKINASE.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sugiki M., Yoshida E., Anai K., Maruyama M.;  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Yano W., Watanabe M.;  
 RT "Oryctolagus cuniculus urokinase-type plasminogen activator, mRNA,  
 RT complete cds.";  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  
 DR EMBL; AY029517; AAK40239.1; -.  
 DR EMBL; AB087224; BAC02685.1; -.  
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
 DR GO; GO:0016301; F:kinase activity; IEA.  
 DR GO; GO:0008233; F:peptidase activity; IEA.  
 DR GO; GO:0004295; F:trypsin activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR009003; Cys\_Ser\_trypsin.

```
DR InterPro: IPR006209; EGF like.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Peptidase_S1.
DR InterPro: IPR001314; Peptidase_S1A.
DR InterPro: IPR008293; Pept_S1A_upA.
DR Pfam: PF00051; kringle; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; Tryp_SPC; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS00070; KRINGLE_2; 1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
DR PIRSF: PIRSF001144; Urk_plasm_act; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48444 MW; 6DD35A371010A6EE CRC64;

Query Match 81.4%; Score 415; DB 6; Length 433;
Best Local Similarity 79.5%; Pred. No. 3.9e-40;
Matches 70; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNCRPN 60
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
70 KTCYHGDSHSGYRGKANTDMDRPLCLAWNSANVLTXTYHAHRPDALQLGLGKHNCRPN 129

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
130 QRRPWCYVQVGLKPLVQECMVHDCSSGK 157

RESULT 4
Q9TVAS PRELIMINARY; PRT; 157 AA.
AC Q9TVAS;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Urokinase plasminogen activator (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=21071388; PubMed=11204721;
RA Balcerzak D., Querengesser L., Dixon W.T., Baracos V.E.;
RT "Coordinate expression of matrix-degrading proteinases and their
activators and inhibitors in bovine skeletal muscle.";
RL J. Anim. Sci. 79:94-107(2001).
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL: AF14761; RAD30301.1; -.
DR HSSP: P00749; LURK.
DR GO: GO:0016301; F:kinase activity; IEA.
DR InterPro: IPR006209; EGF like.
DR InterPro: IPR000001; Kringle.
DR Pfam: PF00051; kringle; 1.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle; 1.
DR SMART: SM00130; KR; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kinase; Kringle.
FT NON_TER 1 157
SQ SEQUENCE 157 AA; 17858 MW; A768D6C72C1F8FB7 CRC64;
```

```
Query Match 77.6%; Score 396; DB 6; Length 157;
Best Local Similarity 76.1%; Pred. No. 2.2e-38;
Matches 67; Conservative 8; Mismatches 13; Indels 0; Gaps 0

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNCRPN 60
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
36 KTCYQNGHSGYRGKANRDLGRCPLAWDSPTVLLKMYHAHRSDAQLGLGKHNCRPN 95

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
96 QRRPWCYVQVGLKPLVQECMVHDCSVGK 123

RESULT 5
Q8C6L2 PRELIMINARY; PRT; 231 AA.
AC Q8C6L2;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Plasminogen activator (Fragment).
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL: AK054349; BAC35743.1; -.
DR F.R. PT0534; PT0534.
DR GO: GO:0004295; F:trypsin activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR009003; Cys Ser trypsin.
DR InterPro: IPR006209; EGF like.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Peptidase_S1.
DR Pfam: PF00051; kringle; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000395; Kringle; 1.
DR SMART: SM00130; KR; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS00070; KRINGLE_2; 1.
FT NON_TER 231 231
SQ SEQUENCE 231 AA; 25510 MW; 25E8980A682737F2 CRC64;

Query Match 75.9%; Score 387; DB 11; Length 231;
Best Local Similarity 72.7%; Pred. No. 3.7e-37;
Matches 64; Conservative 10; Mismatches 14; Indels 0; Gaps 0

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNCRPN 60
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
69 KTCYHGNGDSYRGKANTDKGRPLANAPVILQKPYNAHRPDAISLGLGKHNCRPN 128

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
129 QRRPWCYVQVGLKPLVQECMVHDCSLSK 156

RESULT 6
Q9XT70 PRELIMINARY; PRT; 214 AA.
ID Q9XT70
AC Q9XT70
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
```

01-OCT-2003	(TrEMBLrel. 25, Last annotation update)
DE	Urokinase-type plasminogen activator (Fragment).
OS	Oryctolagus cuniculus (Rabbit).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
NCBI_TaxID=9986;	
[1]	
SEQUENCE FROM N.A.	
RP	TISSUE=Lung;
RC	Yin J., Ideil S.;
RA	"Partial mRNA of rabbit uPA.";
RL	Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC	-1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR	EMBL; AF057647; AAD39351.1; -.
DR	RSSP; P00749; IEDJ.
DR	MEROPS; S01.231; -.
DR	GO; GO:0004363; F:chymotrypsin activity; IEA.
DR	GO; GO:0016301; F:kinase activity; IEA.
DR	GO; GO:0008233; F:peptidase activity; IEA.
DR	GO; GO:0004395; F:trypsin activity; IEA.
DR	GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR	InterPro; IPR009003; Cys_Ser_trypsin.
DR	InterPro; IPR000001; Kringle.
DR	InterPro; IPR001254; Peptidase S1.
DR	InterPro; IPR001314; Peptidase_S1A.
DR	Pfam; PF00051; kringle; 1.
DR	Pfam; PF00089; trypsin; 1.
DR	PRINTS; PRC0722; CHYMOTRYPSIN.
DR	ProDom; PD000395; Kringle; 1.
DR	SMART; SM00130; KR; 1.
DR	SMART; SM00020; Tryp_SPC; 1.
DR	PROSITE; PS00021; KRINGLE_1; 1.
DR	PROSITE; PS00070; KRINGLE_2; 1.
DR	PROSITE; PS00240; TRYPSIN_DOM; 1.
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.
DR	Kw Glycoprotein; Hydrilase; Kinase; Kringle; Protease; Serine protease.
FT	NON TER 214
FT	SEQUENCE 214 AA; 24314 MW; 69975C41C32B0D7E CRC64;
Qy	Query Match 61.4%; Score 313; DB 6; Length 214;
Db	Best Local Similarity 79.3%; Pred.No.1.5e-28;
Db	Matches 54; Conservative 5; Mismatches 10; Indels 0; Gaps 0;
Qy	20 MGRECLPWSGATVLQOTYHAHRSDALQLGLGKHNYCNPDRRPPWCYVQVGLKPLVQEC 79
Db	1 MDRECLAWSANVLTKTYHAHPPDALQLGLGKHNYCNPDRRPPWCYVQVGLKPLVQEC 60
Qy	80 MVHDCADGK 88
Db	61 KVHDSGKK 69
RESULT 7	
Q8MKBI	
ID	Q8MKBI PRELIMINARY; PRT; 564 AA.
AC	Q8MKBI;
DT	01-OCT-2002 (TrEMBLrel. 22, Created)
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE	Tissue-type plasminogen activator.
OS	Oryctolagus cuniculus (Rabbit).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
NCBI_TaxID=9986;	
[1]	
SEQUENCE FROM N.A.	
RA	Sugiki M., Yoshida E., Anai K., Maruyama M.;
PL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC	-1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR	EMBL; AY029518; AAK40240.1; -.
DR	GO; GO:0005576; C:extracellular; IEA.



```
QY 2 TCYEGNGHYRGKASTDTNGRCLPWNATVLQOTYHAHRSDALQLGLGKKNYCRNPDR 61
Db 126 TCYEDQGISYRGTWTAESGAECTNNSSALAKPKYSGRRPDAIRLGLGNHNYCRNPDR 185
QY 62 RRPWCYVQVGLKPLVQECWVHDCADG 87
Db 186 SRPWCYVFKAGKYSSEFCSTPACSEG 211

RESULT 9
Q9BU99 PRELIMINARY; PRT; 516 AA.
AC Q9BU99;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Similar to plasminogen activator, tissue.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RC TISSUE=Skin;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; SC002795; AAH02795.1; -.
DR HSSP; P00750; IASH.
DR GO; GO:0004263; P:chymotrypsin activity; IEA.
DR GO; GO:0008233; P:peptidase activity; IEA.
DR GO; GO:0004295; P:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 516 AA; 57370 MW; BAB31901FDC96800 CRC64;

Query Match 44.3%; Score 226; DB 4; Length 516;
Best Local Similarity 47.7%; Pred. No. 5.8e-18;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRGKASTDTNGRCLPWNATVLQOTYHAHRSDALQLGLGKKNYCRNPDR 61
Db 80 TCYEDQGISYRGTWTAESGAECTNNSSALAKPKYSGRRPDAIRLGLGNHNYCRNPDR 139
QY 62 RRPWCYVQVGLKPLVQECWVHDCADG 87
Db 140 SRPWCYVFKAGKYSSEFCSTPACSEG 165
```

```
RESULT 10
Q86YK8 PRELIMINARY; PRT; 562 AA.
AC Q86YK8;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Tissue plasminogen activator.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Liu Y., Xu L., Zeng Y., He X.;
RT "cDNA of tissue plasminogen activator.";
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY221101; AAC34406.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; P:chymotrypsin activity; IEA.
DR GO; GO:0004295; P:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000083; Fibnctn1.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00058; FNI; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
SQ SEQUENCE 562 AA; 62902 MW; 837D98392F6EDD1F CRC64;

Query Match 44.3%; Score 226; DB 4; Length 562;
Best Local Similarity 47.7%; Pred. No. 6.3e-18;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRGKASTDTNGRCLPWNATVLQOTYHAHRSDALQLGLGKKNYCRNPDR 61
Db 126 TCYEDQGISYRGTWTAESGAECTNNSSALAKPKYSGRRPDAIRLGLGNHNYCRNPDR 185
QY 62 RRPWCYVQVGLKPLVQECWVHDCADG 87
Db 186 SRPWCYVFKAGKYSSEFCSTPACSEG 211

RESULT 11
Q8NG20 PRELIMINARY; PRT; 90 AA.
AC Q8NG20;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Plasminogen/activator kringle.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
```

```
[1]
RN SEQUENCE FROM N.A.
RA Dou D.;
RT "Production of kringle fragment.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF282882; AAM52248.1; -.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kringle.
SQ SEQUENCE 90 AA; 9804 MW; A33887F9FD4C7B1 CRC64;

Query Match 42.8%; Score 218.5; DB 4; Length 90;
Best Local Similarity 49.4%; Pred. No. 6.7e-18;
Matches 41; Conservative 5; Mismatches 36; Indels 1; Gaps 1;

QY 3 CYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDRR 62
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 8 CYFGNGSAYRGTHSLTESGASCLPWN SMILIGKVYTAQNPSAQLGLGKHNYCRNPDDA 67
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 63 RPWCYVQVGLKPLVQECMVHDCA 85
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 68 KWCYTTNPRKLYDCVPPQCA 89
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
Q9BZW1 PRELIMINARY; PRT; 395 AA.
AC Q9BZW1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Neonatal thrombolytic agent alpha-form (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Dou D.;
RT "A brain-type plasminogen activator.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF260825; AAK11956.1; -.
DR HSP; P00750; 1PK2.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00021; KRINGLE_2; 2.
DR PROSITE; PS00240; TRYPSIN DOM; 1.
DR PROSITE; PS00134; TRYPSIN HIS; 1.
DR PROSITE; PS00135; TRYPSIN SER; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
```

KW Serine protease.  
SQ SEQUENCE 562 AA; 63668 MW; F9E6B4C77CB101E8 CRC64;  
Query Match 41.2%; Score 210; DB 6; Length 562;  
Best Local Similarity 43.7%; Pred. No. 4.7e-16;  
Matches 38; Conservative 9; Mismatches 40; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRPCLPWNATVLOQTYHAHRSDALQGLGKHNYCRNPDNR 61  
Db 126 TCYEDQGITYRGTSWTSTESGAEVNNWNTSLASMPYNGRRPDVAKLGLGNHNYCRNPDKD 185

QY 62 RRPWCYVQVGLKPLVOECWVDCADGK 88  
Db 186 SKPWCYIFKAEKYPDFCSTACTKEK 212

RESULT 14  
ID Q8KOD2 PRELIMINARY; PRT; 517 AA.  
AC Q8KOD2;  
DT 01-OCT-2002 (Tremblrel. 22, Created)  
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC Tissue=Kidney;  
RA Strausberg R.;  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  
DR EMBL; BC031775; AAH31775.1; -  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0005508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR009003; Cys Ser trypsin.  
DR InterPro; IPR000742; EGF 2.  
DR InterPro; IPR001881; EGF like.  
DR InterPro; IPR006209; EGF like.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR Pfam; PF00008; EGF; 2.  
DR Pfam; PF00051; Kringle; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00181; EGF; 3.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS00022; EGF 1; 3.  
DR PROSITE; PS01186; EGF 2; 2.  
DR PROSITE; PS00021; KRINGLE 1; 1.  
DR PROSITE; PS00070; KRINGLE 2; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;  
KW Kringle; Protease; Serine protease.  
SQ SEQUENCE 517 AA; 57326 MW; 3855A42035A5EA59 CRC64;

Query Match 41.1%; Score 209.5; DB 11; Length 517;  
Best Local Similarity 45.8%; Pred. No. 4.9e-16;  
Matches 38; Conservative 13; Mismatches 31; Indels 1; Gaps 1;

QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNATVLOQTYHAHRSDALQGLGKHNYCRNPDNR 62  
Db 151 CVYGDGYSYRGKSVKVNQNPCLYNNSHLLQETNNMFEDAEHTGIAHNFRCRNPDDGH 210

QY 63 RRPWCYVQVGLKPLVOECWVDCADGK 84  
Db 211 KPCFCYKVNSEKVKWYCDVTV 233

RESULT 15  
ID Q800Y7 PRELIMINARY; PRT; 540 AA.  
AC Q800Y7;  
DT 01-JUN-2003 (Tremblrel. 24, Created)  
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Hepatocyte growth factor activator (Fragment).  
OS Meleagris gallopavo (Common turkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Meleagris.  
OX NCBI\_TaxID=9103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=22122796; PubMed=12128063;  
RA Holsberger D.R., Becker A.E., Thurston R.J., Rice C.D.;  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
EMBL; AV216598; AAO46038.1; -  
DR GO; GO:0005576; C:extracellular; IEA.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR009003; Cys Ser trypsin.  
DR InterPro; IPR000742; EGF 2.  
DR InterPro; IPR006209; EGF like.  
DR InterPro; IPR000083; Fibnctnl.  
DR InterPro; IPR000562; FN\_Type\_11.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00008; EGF; 2.  
DR Pfam; PF00051; Kringle; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00181; EGF; 2.  
DR SMART; SM00058; FN1; 1.  
DR SMART; SM00059; FN2; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS00022; EGF 1; 2.  
DR PROSITE; PS01186; EGF 2; 1.  
DR PROSITE; PS01253; FIBRONECTIN 1; 1.  
DR PROSITE; PS00021; KRINGLE 1; 1.  
DR PROSITE; PS00070; KRINGLE 2; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
FT NON\_TER 1 540  
FT NON\_TER 540 540  
SQ SEQUENCE 540 AA; 61159 MW; 0BB3B4A89C0B577F CRC64;

Query Match 39.8%; Score 202; DB 13; Length 540;  
Best Local Similarity 43.4%; Pred. No. 3.9e-15;  
Matches 36; Conservative 10; Mismatches 37; Indels 0; Gaps 0

```
QY 3 CYEGNGHYRGKASTDTWGRPCLPWNSATVLTQTYHAHRSALQGLGKHYCRNPDNR 62
Db 179 CYRNGTEYRGKASTDTWGRPCLPWNSATVLTQTYHAHRSALQGLGKHYCRNPDNR 238
QY 63 RPWCYVQVGLKPLVQECMVHDC 85
Db 239 KPCWYIMKXNSLSWEYCNITSCA 261

RESULT 16
Q14520 PRELIMINARY; PRT; 560 AA.
AC Q14520; 000663;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE HGF activator like protein (Hyaluronan binding protein 2).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1] : : : : :
RP SEQUENCE FROM N.A.
RA Kitamura N.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96425001; PubMed=8827452;
RA Choi-Mura N.H.; Tobe T.; Sumiya J.; Nakano Y.; Sano Y.; Mada T.;
RA Tomita M.;
RT "Purification and characterization of a novel hyaluronan-binding
RT protein (PHBP) from human plasma: it has three EGF, a kringle and a
RT serine protease domain, similar to hepatocyte growth factor
RT activator.";
RL J. Biochem. 119:1157-1165 (1996).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon, and Kidney;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; D49742; BAA08576.1; -
DR EMBL; S83182; ABA46909.1; -
DR EMBL; BC031412; AAH31412.1; -
DR PIR; JC4795; JC4795.
DR HSSP; P00763; LDPO.
DR MEROPS; S01.033; -
DR Genew; HGNC:4798; HARP2.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0005533; F:glycosaminoglycan binding; TAS.
DR GO; GO:0007155; P:cell adhesion; TAS.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYNOTRYPsin.
DR PRINTS; PRO0018; KRINGLE.
DR PRODom; PD000395; Kringle; 1.
DR PROSITE; PS00022; EGF 1; 3.
DR PROSITE; PS01186; EGF 2; 2.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
DR PROSITE; PS00240; TRYPSIN DOM; 1.
DR PROSITE; PS00134; TRYPSIN HIS; 1.
DR PROSITE; PS00135; TRYPSIN_S1; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 560 AA; 62671 MW; 5C1907230784ACD4 CRC64;
```

```
Query Match 37.5%; Score 193.5; DB 4; Length 560;
Best Local Similarity 41.7%; Pred. No. 4e-14;
Matches 35; Conservative 15; Mismatches 33; Indels 1; Gaps 1

QY 3 CYEGNGHYRGKASTDTWGRPCLPWNSATVLTQTYHAHRSALQGLGKHYCRNPDNR 62
Db 194 CYVGDGYSYRGKASTDTWGRPCLPWNSATVLTQTYHAHRSALQGLGKHYCRNPDNR 253
QY 63 RPWCYVQVGLKPLVQECMVHDC 85
Db 254 KPCWYIMKXNSLSWEYCNITSCA 277

RESULT 17
Q90675 PRELIMINARY; PRT; 202 AA.
AC Q90675;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE tissue-type plasminogen activator (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White Leghorn;
RX MEDLINE=97199025; PubMed=9047000;
RA Johnson A.L.; Bridgham J.T.; Anthony R.V.;
RT "Expression of avian urokinase and tissue-type plasminogen activator
RT messenger ribonucleic acid during follicle development and atresia.";
RL Biol. Reprod. 56:581-588 (1997).
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; U31988; AAA74955.1; -
DR HSSP; P00750; 1RTF.
DR MEROPS; S01.232; -
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0018; KRINGLE.
DR PRODom; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 2.
DR PROSITE; PS00240; TRYPSIN DOM; 1.
DR PROSITE; PS00134; TRYPSIN HIS; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
FT NON_TER 1
FT NON_TER 202
SQ SEQUENCE 202 AA; 23148 MW; 049DB42941D83AFC CRC64;

Query Match 37.5%; Score 191; DB 13; Length 202;
Best Local Similarity 43.9%; Pred. No. 2.6e-14;
Matches 36; Conservative 7; Mismatches 39; Indels 0; Gaps 0

QY 3 CYEGNGHYRGKASTDTWGRPCLPWNSATVLTQTYHAHRSALQGLGKHYCRNPDNR 62
Db 41 CYTNGGLAYRGKASTDTWGRPCLPWNSATVLTQTYHAHRSALQGLGKHYCRNPDNR 100
QY 63 RPWCYVQVGLKPLVQECMVHDC 84
Db 101 QPWCYVQVGLKPLVQECMVHDC 122
```

```

RESULT 18
QSVCS4      QSVCS4      PRELIMINARY;      PRT;      653 AA.
AC          QSVCS4;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Liver;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBSJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR HSSP; P00761; IAN1.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004235; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR00742; EGF_2.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR008210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00053; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; ENTYPPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS01253; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00021; KRINGLE_2; 1.
DR PROSITE; PS00070; TRYPSIN_DOM; 1.
DR PROSITE; PS00240; TRYPSIN_HIS; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
KW Kringle; Protease; Serine protease.
SQ SEQUENCE 653 AA; 70553 MW; FE18D90174ED6FDD CRC64;

Query Match      36.9%; Score 188; DB 11; Length 653;
Best Local Similarity 53.0%; Pred. No. 2e-13;
Matches 35; Conservative 6; Mismatches 25; Indels 0; Gaps 0;

QY 3 CYEGNGHYRGKASTDTMGRCLPWN SATVLQOITYHAHRS DALQLGLGKKNYCNPNRR 62
Db 283 CFLNGTGYRGVASTAASGLSCLAWNSDLQYELHVDVSAVAALLGLGPHAYCRNPKDE 342
QY 63 RPKCVY 68
Db 343 RPKCVY 348

```

```

RESULT 19
O35727      O35727      PRELIMINARY;      PRT;      597 AA.
AC          O35727;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Factor XII.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Liver;
RA Schloesser M.; Schwager S.; Engel W.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBSJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; X9571; CAAG7891.1; -.
DR HSSP; P00760; 1AQ7.
DR MEROPS; S01.211; -.
DR MGD; MGI:1891012; F12.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004235; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; ENTYPPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS01253; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00021; KRINGLE_2; 1.
DR PROSITE; PS00070; TRYPSIN_DOM; 1.
DR PROSITE; PS00240; TRYPSIN_HIS; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 597 AA; 65638 MW; F3AC07C37D0C0FBA CRC64;

Query Match      36.8%; Score 186.5; DB 11; Length 597;
Best Local Similarity 44.0%; Pred. No. 2.8e-13;
Matches 37; Conservative 8; Mismatches 34; Indels 5; Gaps 2

QY 2 TCYEGNGHYRGKASTDTMGRCLPWN SATVLQOITYHAHRS DALQLGLGKKNYCNPNR 60
Db 283 CFLNGTGYRGVASTAASGLSCLAWNSDLQYELHVDVSAVAALLGLGPHAYCRNPKDE 342
QY 63 RPKCVY 68
Db 343 RPKCVY 348

```



Best Local Similarity 44.6%; Pred. No. 3.2e-12;  
Matches 37; Conservative 6; Mismatches 35; Indels 5; Gaps 2;  
QY 3 CYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOQTY-HAHRSDALQGLGKHNCRPNDR 61  
Db 217 CYSRDLGSLYRGVQAQTTLSGAPCPQWAS----EATYNNVTAQALNWGLGDAFCRPNPD 272  
QY 62 RRPWCYVQUGLKLPLVQSCWHDHC 84  
Db 273 TRPWCFCVWRGDLQSLWQYCLRLARC 295

RESULT 22  
Q8AXX3 PRELIMINARY; PRT; 421 AA.  
AC Q8AXX3 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-WAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Kremen2.  
GN KRM2.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
OC Xenopodinae; Xenopus.  
OC NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22308873; PubMed=12421700;  
RA Davidson G., Mao B., Del Barco Barrantes I., Niehrs C.;  
RT "Kremen proteins interact with Dickkopf1 to regulate anteroposterior  
CNS patterning";  
RL Development 129:5587-5596(2002).  
DR EMBL; AY150813; AAN64661.1; -.  
DR InterPro; IPR000859; CUB.  
DR InterPro; IPR000001; Kringle.  
DR Pfam; PF00431; CUB; 1.  
DR Pfam; PF00051; Kringle; 1.  
DR Pfam; PF01822; WSC; 1.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00042; CUB; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00321; WSC; 1.  
DR PROSITE; PS01180; CUB; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS00070; KRINGLE\_2; 1.  
SQ SEQUENCE 421 AA; 46274 MW; 74D04DB1682CD15F CRC64;

Query Match 33.8%; Score 172.5; DB 13; Length 421;  
Best Local Similarity 50.7%; Pred. No. 8.3e-12;  
Matches 34; Conservative 4; Mismatches 26; Indels 3; Gaps 2;  
QY 3 CYEGNGHFGYRGKAS-TDTMGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNCRPNDR 61  
Db 29 CFTVGRDYGTVSQAGPGTGLVWQTT--CHLYNAQSDPDGELGLGNHCRPNPD 86  
QY 62 RRPWCYV 68  
Db 87 VQPCYV 93

RESULT 23  
Q81Z25 PRELIMINARY; PRT; 615 AA.  
AC Q81Z25 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-WAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Coagulation factor XII-Mie.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wada H., Nishio K., Nakatani K., Kasai Y., Abe Y., Nobori T.;  
RT "Molecular characterization of coagulation factor XII-Mie";  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB095845; BAC23095.1; -.  
DR GO; GO:0005576; C:extracellular; IEA.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR000083; Fibrinogen.  
DR InterPro; IPR000562; FN Type II.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR001254; Peptidase S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00008; EGF; 2.  
DR Pfam; PF00039; fn1; 1.  
DR Pfam; PF00040; fn2; 1.  
DR Pfam; PF00051; kringle; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00013; FNTYPEII.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000395; FN Type II; 1.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00181; EGF; 2.  
DR SMART; SM00179; EGF\_Ca; 2.  
DR SMART; SM00058; FN1; 1.  
DR SMART; SM00059; FN2; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS00022; EGF\_1; 2.  
DR PROSITE; PS01186; EGF\_2; 1.  
DR PROSITE; PS01253; FIBRONECTIN\_1; 1.  
DR PROSITE; PS00023; FIBRONECTIN\_2; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS00070; KRINGLE\_2; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00021; TRYPSIN\_SER; 1.  
SQ SEQUENCE 615 AA; 67735 MW; 030508870A0C7EDB CRC64;

Query Match 33.4%; Score 170.5; DB 4; Length 615;  
Best Local Similarity 48.5%; Pred. No. 2.1e-11;  
Matches 33; Conservative 6; Mismatches 24; Indels 5; Gaps 2;  
QY 2 TCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOQTY-HAHRSDALQGLGKHNCRPNDR 60  
Db 216 SCVDGRLSVRGLARTLSGAPCPQWAS----EATYNNVTAQALNWGLGDAFCRPNDR 271  
QY 61 RRPWCYV 68  
Db 272 DIRPWCYV 279

RESULT 24  
Q8AXY6 PRELIMINARY; PRT; 947 AA.  
AC Q8AXY6 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Muscle-specific receptor tyrosine kinase MusK.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.  
OX NCBI\_TaxID=9031;  
RN SEQUENCE FROM N.A.  
RX MEDLINE=20538710; PubMed=11083926;  
RA Ip F.C., Glass D.G., Gies D.R., Cheung J., Lai K.O., Fu A.K.,  
RA Yancopoulos G.D., Ip N.Y.;  
RT "Cloning and characterization of muscle-specific kinase in chicken.";  
RL Mol. Cell. Neurosci. 16:661-673(2000).  
RN [2]  
RN SEQUENCE FROM N.A.  
RA Gies D., Glass D.J., Yancopoulos G.D.;  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY143173; AAN05008.1; -  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005874; C:microtubule; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
DR GO; GO:0005199; F:structural molecule activity; IEA.  
DR GO; GO:0004888; F:transmembrane receptor activity; IEA.  
DR GO; GO:0007275; P:development; IEA.  
DR GO; GO:0007018; P:microtubule-based movement; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR002453; Beta\_tubulin.  
DR InterPro; IPR000024; F2\_domain.  
DR InterPro; IPR0003599; IG.  
DR InterPro; IPR007110; IG-like.  
DR InterPro; IPR0003598; IG\_c2.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR0000719; Prot\_kinase.  
DR InterPro; IPR002290; Ser\_thr\_kinase.  
DR InterPro; IPR001245; Tyr\_kinase.  
DR InterPro; IPR008266; Tyr\_kinase\_AS.  
DR Pfam; PF01392; Fz; 1.  
DR Pfam; PF00047; Ig; 3.  
DR Pfam; PF00051; kringle; 1.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00018; KRINGLE.  
DR PRINTS; PR00189; TYRKINASE.  
DR ProDom; PD000395; IG LIKE; 3.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR SMART; SM00409; IG; 3.  
DR SMART; SM00408; Igc2; 3.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00220; S\_TKc; 1.  
DR SMART; SM00219; TYRKc; 1.  
DR PROSITE; PS00038; FZ; 1.  
DR PROSITE; PS00835; IG LIKE; 3.  
DR PROSITE; PS00021; KRINGLE 1; 1.  
DR PROSITE; PS00070; KRINGLE 2; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS00228; TUBULIN\_B\_AUTOREG; 1.  
KW Receptor; Kinase.  
SQ SEQUENCE 947 AA; 105588 MW; E6C3FCC5796CC8B CRC64;  
  
Query Match 32.1%; Score 163.5; DB 13; Length 947;  
Best Local Similarity 38.8%; Pred. No. 2.2e-10;  
Matches 38; Conservative 6; Mismatches 27; Indels 27; Gaps 6;  
  
QY 1 KTCYENGHFYRGKASTDTMGRLPCLPWSATVLQTYHAHR-----SDALQLGLGK 51  
Db 463 RTCYSGNGFYQGWANVTASGIPCKQWS-----DQAPLHRRTPQVFPFLSDA----- 510  
  
QY 52 HNYCRNP-DNRREPCVQVGLKPLV--QECMVHDCAD 86  
Db 511 ENYCRNPGGENRPECWYTK---DPSVTWEYCSVSPCGD 545  
  
Query Match 32.1%; Score 163.5; DB 13; Length 947;  
Best Local Similarity 38.8%; Pred. No. 2.2e-10;  
Matches 38; Conservative 6; Mismatches 27; Indels 27; Gaps 6;

ID Q46506 PRELIMINARY; PRT; 454 AA.  
AC Q46506;  
DT 01-JUN-1998 (TREMblrel. 06, Created)  
DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)  
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)  
DE Apolipoprotein a (Fragment).  
GN BAAPOA.  
OS Papio hamadryas (Hamadryas baboon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecoidea; Papio.  
OX NCBI\_TaxID=9557;  
RN [1]  
RN SEQUENCE FROM N.A.  
RA Cox L.A., Jett C., Hixson J.E.;  
RT "Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice Site Mutation is Associated with Deletion of a Single Exon in a Null Allele.";  
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
CC -/- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -/- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.  
DR EMBL; AF029691; AAB97886.1; -  
DR HSP; P00747; 2PK4.  
DR MEROPS; S01.999; -  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00051; kringle; 2.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000395; Kringle; 2.  
DR SMART; SM00130; KR; 2.  
DR SMART; SM00020; Tryp\_Spc; 1.  
DR PROSITE; PS00021; KRINGLE 1; 2.  
DR PROSITE; PS00070; KRINGLE 2; 2.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Glycoprotein; Hydrolase; Kringle; Lipoprotein; Protease;  
KW Serine protease.  
FT NON\_TER 1  
SQ SEQUENCE 454 AA; 50041 MW; 974E3074C187B2F CRC64;  
  
Query Match 31.0%; Score 158; DB 6; Length 454;  
Best Local Similarity 39.8%; Pred. No. 4.4e-10;  
Matches 35; Conservative 10; Mismatches 31; Indels 12; Gaps 4;  
  
QY 3 CYEGNGHFYRGKASTDTMGRLPCLPWSATVLQ--QTYHAHRSDALQLGLGKHYCRNPDN 60  
Db 102 CYHGDGQSYRGFSFTTGTCTQSSSMTPHQHKRTPEHPNDGLTM-----NYCRNPD 156  
  
QY 61 RRRPCVQVGLKPLV--QECMVHDCAD 86  
Db 157 DTGFWCFET---MDPSVRWEYCNLTRCSD 181  
  
RESULT 26  
Q9UIRS PRELIMINARY; PRT; 113 AA.  
ID Q9UIRS  
AC Q9UIRS;  
DT 01-MAY-2000 (TREMblrel. 13, Created)  
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
DE Apolipoprotein(a) (Fragment).  
GN APOA.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.



```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21181705; PubMed=11285247;
RA Ogorelkova M., Kraft H.G., Enholm C., Utermann G.;
RT "Single nucleotide polymorphisms in exons of the apo(a) kringle IV
RT types 6 to 10 domain affect lip(a) plasma concentrations and have
RT different patterns in Africans and Caucasians.";
RL Hum. Mol. Genet. 10:815-824 (2001).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF158663; AAF03680.1; -.
DR HSSP; P00747; 1PWK.
DR InterPro; IPR000001; Kringle.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kringle; Lipoprotein.
FT NON_TER 1
FT NON_TER 113
SQ SEQUENCE 113 AA; 12685 MW; F3D65681D985253A CRC64;

Query Match 30.6%; Score 156; DB 4; Length 113;
Best Local Similarity 37.8%; Pred. No. 1.7e-10;
Matches 34; Conservative 12; Mismatches 32; Indels 12; Gaps 4;

Qy 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATV--LQOTYHAHSDALQLGLGKHNYCRNP 59
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
9 RCQYHGNGSQYRGFTSVTGTCTQSSMTPHRHQRTPEYFNDGLTM----NVCNRP 63

Qy 59 DNRERPQYVQVGLKPLV--QCVNHDCA 86
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
64 DADTGPWCFT---MDPSIRWEYCNLTTCSD 90

RESULT 27
Q7SXB3 PRELIMINARY; PRT; 263 AA.
AC Q7SXB3
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=AB; TISSUE=Body;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Atschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Iqbalan N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalak U., Smallus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human

```

```

RT and mouse cDNA sequences.";
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=AB; TISSUE=Body;
RA Strausberg R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC055675; AAHS5675.1; -.
KW Hypothetical protein.
SQ SEQUENCE 263 AA; 28777 MW; 8B8EC117BC7C9A58 CRC64;

Query Match 30.5%; Score 155.5; DB 13; Length 263;
Best Local Similarity 35.6%; Pred. No. 4.8e-10;
Matches 31; Conservative 8; Mismatches 39; Indels 9; Gaps 2

Qy 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLTQTYHAHSDALQLGLGKHNYCRNPDN 60
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
23 KDCITNGEDYRGTOQTQKTSSTGSLTSLWSRLN-----KFKSQTGVGDRHFCRNPDG 74

Qy 61 RRRPWCYVQVGLKPLVQ--CMVHDCAD 86
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
75 SNKPWCYVSGSGGETKKEZACDIRICQD 101

RESULT 28
O18783 PRELIMINARY; PRT; 806 AA.
AC O18783
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Plasminogen.
OS Macropus eugenii (Tamar wallaby).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
OX NCBI_TaxID=9315;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Liver;
RA MEDLINE=98004511; PubMed=9342350;
RA Lawn R.M., Schwartz K., Patthy L.;
RT "Convergent evolution of apolipoprotein(a) in primates and hedgehog.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:11992-11997 (1997).
CC -!- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
DR EMBL; AF012297; AAB65760.1; -.
DR HSSP; P00747; 5HEG.
DR MEROPS; S01.233; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003809; F:thrombin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0007586; P:blood coagulation; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan app.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR003566; Peptidase_S1A_pr.
DR Pfam; PF00051; Kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR PRODOM; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 5.
DR PROSITE; PS00070; KRINGLE_2; 5.

```

DR PROSITE; PS00240; TRYPSIN DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.  
SQ SEQUENCE 806 AA; 90981 MW; 95FAA86DC20064D5 CRC64;  
  
Query Match  
Best Local Similarity 30.2%; Score 154; DB 6; Length 806;  
Matches 33; Conservative 11; Mismatches 25; Indels 24; Gaps 5;  
  
QY 3 CYEGNGHYRGKASTDTMGRCPLPNSATVLCQYTHAHR-----SDALQLGLGKHN 54  
DB 371 CYEGNGHYRGKASTDTMGRCPLPNSATVLCQYTHAHR-----SDALQLGLGKHN 54  
QY 55 CRNPDRRPPWCYVQVGLKPLV--QECMVHDC 85  
DB 420 CRNPDRRPPWCYVQVGLKPLV--QECMVHDC 85  
  
RESULT 29  
QY 09UIR7 PRELIMINARY; PRT; 113 AA.  
AC Q9UIR7;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DE 01-MAY-2003 (TrEMBLrel. 23, Last annotation update)  
DE Apolipoprotein(a) (Fragment).  
GN APOA.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
SQ SEQUENCE FROM N.A.  
RX MEDLINE=21181705; PubMed=11285247;  
RA Ogorekova M., Kraft H.G., Ehnholm C., Utermann G.;  
RT "Single nucleotide polymorphisms in exons of the apo(a) kringles IV  
RT types 6 to 10 domain affect Lp(a) plasma concentrations and have  
RT different patterns in Africans and Caucasians."  
RL Hum. Mol. Genet. 10:815-824 (2001).  
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  
DR EMBL; AF158659; AAF03678.1; -.  
DR EMBL; AF158658; AAF03678.1; JOINED.  
DR HSSP; P00747; 2PK4.  
DR InterPro; IPR000001; Kringle.  
DR Pfam; PF00051; kringle; 1.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00130; KR; 1.  
DR PROSITE; PS00021; KRINGLE 1; 1.  
DR PROSITE; PS00070; KRINGLE 2; 1.  
KW Glycoprotein; Kringle; Lipoprotein.  
FT NON\_TER 1  
FT NON\_TER 113  
SQ SEQUENCE 113 AA; 12815 MW; 4F80ADF8708548CB CRC64;  
  
Query Match  
Best Local Similarity 30.0%; Score 153; DB 4; Length 113;  
Matches 33; Conservative 8; Mismatches 35; Indels 8; Gaps 3;  
  
QY 3 CYEGNGHYRGKASTDTMGRCPLPNSATVLCQYTHAHRSDALQLGLGKHNCRPNRR 62  
DB 11 CYHGNGQSYRGTSFTVTGRTQCSWSMT---PHWHQRTTEYPNGGLTRNCRNDAEI 67  
QY 63 RWCYVQVGLKPLV--QECMVHDC 84  
DB 68 RWCYVQVGLKPLV--QECMVHDC 84  
  
RESULT 30  
QY 08WMR1 PRELIMINARY; PRT; 359 AA.  
ID Q8WMR1  
AC Q8WMR1;

DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Plasminogen (Fragment).  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
SQ SEQUENCE FROM N.A.  
RA Pirie-Shepherd S.R., Coffman K.T., Resnick D., Chan R., Kisker O.,  
RA Folkman J., Waters D.J.;  
RT "Angiostatin is Detectable in the Urine of Dogs with Spontaneous Bone  
RT Cancer."  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.  
DR EMBL; AY069985; AAL5819.1; -.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0003809; F:thrombin activity; IEA.  
DR GO; GO:0007596; P:blood coagulation; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR003966; Peptidase\_S1A\_pr.  
DR Pfam; PF00051; Kringle; 4.  
DR PRINTS; PR00018; KRINGLE.  
DR PRINTS; PR01505; PROTHROMBIN.  
DR ProDom; PD000395; Kringle; 4.  
DR SMART; SM00130; KR; 4.  
DR PROSITE; PS00021; KRINGLE 1; 4.  
DR PROSITE; PS00070; KRINGLE 2; 4.  
KW Glycoprotein; Kringle.  
FT NON\_TER 1  
FT NON\_TER 359  
SQ SEQUENCE 359 AA; 41172 MW; 776D33F4AB0BDD9E CRC64;  
  
Query Match  
Best Local Similarity 29.7%; Score 151.5; DB 6; Length 359;  
Matches 31; Conservative 5; Mismatches 21; Indels 13; Gaps 3  
  
QY 3 CYEGNGHYRGKASTDTMGRCPLPNSATVLCQYTHAHRSDALQLGLGKHNCRN 57  
DB 282 CYHGNGQSYRGTSFTVTGRTQCSWSMT-----PHRHEKTPHFPEAGL-TWNYCRN 333  
QY 58 PDNRREPWCY 67  
DB 334 PDNRREPWCY 343  
  
RESULT 31  
QY 09ROW3 PRELIMINARY; PRT; 812 AA.  
AC Q9ROW3;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Plasminogen protein precursor (EC 3.4.21.7).  
GN PLASMINOGEN.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
SQ SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Bangert K., Johnsen A.H., Thorsen S.;  
RT "Rat plasminogen: cDNA and gene structure."  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
SQ SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=91350378; PubMed=1645711;  
RA Kanalas J.J., Makker S.P.;  
RT "Identification of the rat Heymann nephritis autoantigen (GP330) as a



[illegible]

```
DR PRINTS; PR000395; TYRKINASE..
DR ProDom; PD000395; Kringle; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00338; FZ; 1.
DR PROSITE; PS00835; IG LIKE; 1.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00021; KRINGLE 2; 1.
DR PROSITE; PS00070; KRINGLE 1; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
SQ SEQUENCE 944 AA; 105053 MW; 1EBD416BE3170401 CRC64;

Query Match 29.1%; Score 148.5; DB 11; Length 944;
Best Local Similarity 41.8%; Pred. No. 1.3e-09;
Matches 38; Conservative 9; Mismatches 31; Indels 13; Gaps 7;

QY 3 CYEGNGHFYRGKASTDTMGRCPLPNSATVLQTYHAHR---SDALQGLGKKNYCNPD 59
DB 316 CYNGSGADYRGMASTTSKHQCQW---ALQHP-HSHRUSSTFEPLG-GGHAYCRNPG 369
QY 60 NRRR-PWCYVQVGLKPLVQECMVHDC--DG 87
DB 370 QMGEGPWCFTQ-NKNRVAVELCDVPFCSPRDG 399

RESULT 36
Q9UIR8 PRELIMINARY; PRT; 105 AA.
AC Q9UIR8
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAY-2003 (Tremblrel. 23, Last annotation update)
DE Apolipoprotein(a) (Fragment).
GN APOA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21181705; PubMed=11285247;
RA Ogorelkova M., Kraft H.G., Ehnholm C., Utermann G.;
RT "Single nucleotide polymorphisms in exons of the apo(a) kringles IV
RT types 6 to 10 domain affect Lp(a) plasma concentrations and have
RT different patterns in Africans and Caucasians.";
RL Hum. Mol. Genet. 10:815-824(2001).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF158656; AAF03677.1; -.
DR EMBL; AF158660; AAF03679.1; JOINED.
DR HSSP; P00747; 2PK4.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
KW Glycoprotein; Kringle; Lipoprotein.
FT NON_TER 1
FT NON_TER 113
SQ SEQUENCE 113 AA; 12697 MW; 51D461D9C66312E CRC64;

Query Match 28.8%; Score 147; DB 4; Length 113;
Best Local Similarity 36.0%; Pred. No. 1.9e-09;
Matches 31; Conservative 14; Mismatches 33; Indels 8; Gaps 4.

QY 3 CYEGNGHFYRGKASTDTMGRCPLPNSATVLQTYHAHRSDALQGLGKKNYCNPDNR 62
DB 11 CYHGDGSRYSFSTTGTGTCQSSWT---PHHWQRTTEYPNGGLTRNYCRNPDAEI 67

QY 63 RWCYVQVGLKPLV--QECMVHDCAD 86
DB 68 QPWCYT---MDPCVRWEYCNLTQCSE 90

RESULT 38
Q9HIV4 PRELIMINARY; PRT; 648 AA.
AC Q9HIV4
DT 01-MAY-2000 (Tremblrel. 16, Created)
DT 01-MAY-2001 (Tremblrel. 16, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE DJ1182A14.3 (Similar to MST1 (Macrophage stimulating 1 (Hepatocyte
DE growth factor-like))).
GN DJ1182A14.3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX Bird C.;
```

```

Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
EMBL; AL137798; CAC17639.1; -.
DR HSSP; P00747; SHPG.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan app.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00051; kringle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; K:kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 3.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
SQ SEQUENCE 648 AA; 72781 MW; 4CE077057350E463 CRC64;

Query Match 28.8%; Score 147; DB 4; Length 648;
Best Local Similarity 33.0%; Pred. No. 1.3e-08;
Matches 32; Conservative 13; Mismatches 36; Indels 16; Gaps 4;

QY 2 TCYENGHFYRGKASTDTMGRCPLPWNSTATVLCQTYHAHRSALQLGLG--KHNCRNPD 59
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 237 SCFRKGSGYRGATNTTAGVPCQRWDA-----QIPHCHRTPEKYACKDLRFNCFNPD 291

QY 60 MRRRPWCY-----VQVGLKPLVQEC-----MVHDCADG 87
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 292 GSEAPWCFTLRPGKRVGFCYQIRRCTDDVRPQDCVHG 328

RESULT 39
Q9BRB6 PRELIMINARY; PRT; 393 AA.
AC Q9BRB6;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Similar to receptor tyrosine kinase-like orphan receptor 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
EMBL; BC006374; AA06374.1; -.
DR HSSP; P00747; ICEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0004888; F:transmembrane receptor activity; IEA.
DR GO; GO:0007275; P:development; IEA.
DR InterPro; IPR000024; Fz_domain.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF00047; Ig; 1.

Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
EMBL; AL137798; CAC17639.1; -.
DR HSSP; P00747; SHPG.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan app.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00051; kringle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; K:kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 3.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
SQ SEQUENCE 648 AA; 72781 MW; 4CE077057350E463 CRC64;

Query Match 28.6%; Score 146; DB 4; Length 393;
Best Local Similarity 44.1%; Pred. No. 9.6e-09;
Matches 30; Conservative 5; Mismatches 25; Indels 8; Gaps 3;

QY 3 CYENGHFYRGKASTDTMGRCPLPWNSTATVLCQTYHAHRSALQLGLG--LGKHNCRNPD 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 313 CYNSTGVDRGTGVTSGRCQCPWNS-----QVPHHTFTALRPPELNGGHSYCRNPGN 367

QY 61 RRR-PWCY 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 368 QKEAPWCF 375

RESULT 40
Q13208 PRELIMINARY; PRT; 567 AA.
AC Q13208;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hepatocyte growth factor-like protein homolog.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20191171; PubMed=10728827;
RA Degen S.J.F., McDowell S.A., Waltz S.E., Gould F., Stuart L.A.,
RA Carritt B.;
RT "Structure of the human D1F15S1A locus: a chromosome 1 locus with 97%
RT identity to the chromosome 3 gene coding for hepatocyte growth factor-
RT like protein.";
RL DNA Seq. 8:409-413 (1998).
CC -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
EMBL; U28054; AAC63092.1; -.
DR HSSP; P00747; 2PK4.
DR MEROPS; S01.977; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003809; F:thrombin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0007596; P:blood coagulation; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR003966; Peptidase_S1A_pr.
DR Pfam; PF00051; kringle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; K:kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 3.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
```

```

SQ SEQUENCE 567 AA; 64117 MW; 3FC38B07F1645810 CRC64;
Query Match 28.6%; Score 146; DB 4; Length 567;
Best Local Similarity 33.0%; Pred. No. 1.4e-08;
Matches 32; Conservative 12; Mismatches 37; Indels 16; Gaps 4;

QY 2 TCYEGNGHFYRGKASTDTMGRCPLPWNASATVQLQTYHAHRSALQIGL--KHNYCRNPD 59
Db 257 SCFRGEGYRGVGTANTTTAGVPCQRDA-----QIPHOHRFTPEKYACKDLRNEFCRNPD 311
QY 60 NRRRPWCY-----VQVGLKPLVQEC-----MVHDCADG 87
Db 312 GSEAPWCFTLRPGTRVGFQVQIRRCTDDVDPQDCYHG 348

RESULT 41
Q8BNP9 ID Q8BNP9 PRELIMINARY; PRT; 937 AA.
AC Q8BNP9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Receptor tyrosine kinase-like orphan receptor 1.
GN ROR1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=adipose tissue;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK040896; BAC38068.1; -
DR MGD; MGI:1347520; Ror1.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004574; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0004888; F:transmembrane receptor activity; IEA.
DR GO; GO:0007275; P:development; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000024; Fz domain.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig C2.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000395; Kringle; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00220; S_TK; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS50038; FZ; 1.
DR PROSITE; PS50038; FZ; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
SQ SEQUENCE 937 AA; 104118 MW; CB440323CA66EC28 CRC64;
Query Match 28.6%; Score 146; DB 11; Length 937;
Best Local Similarity 44.4%; Pred. No. 2.4e-08;
Matches 30; Conservative 5; Mismatches 25; Indels 8; Gaps 3;

QY 3 CYEGNGHFYRGKASTDTMGRCPLPWNASATVQLQTYHAHRSALQIGL--LKHNYCRNPD 60
Db 313 CYNSTGVDRGTVSVTKSGRCQCPWNS-----QYPHTSTALRFFELNGHGHVYCRNPGN 367
QY 61 RRR-PWCY 67
Db 368 QKEAPWCF 375

RESULT 42
Q8BG10 ID Q8BG10 PRELIMINARY; PRT; 937 AA.
AC Q8BG10;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Receptor tyrosine kinase-like orphan receptor 1.
GN ROR1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=adipose tissue;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK046699; BAC32840.1; -
DR MGD; MGI:1347520; Ror1.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004574; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0004888; F:transmembrane receptor activity; IEA.
DR GO; GO:0007275; P:development; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000024; Fz domain.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig C2.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000395; Kringle; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00220; S_TK; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS50038; FZ; 1.
DR PROSITE; PS50038; FZ; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

```

```
DR PROSITE; PS50070; KRINGLE 2; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
SQ SEQUENCE 937 AA; 104088 MW; D6F2D84B67D03C69 CRC64;

Query Match
Best Local Similarity 28.6%; Score 146; DB 11; Length 937;
Matches 30; Conservative 5; Mismatches 25; Indels 8; Gaps 3;

QY 3 CYEGNGHFRGKASTDTMGRPCLPWNATVLTQTYHAHRSDALQLG--LGHKHYCRNPDN 60
DB 313 CYNSTGVDRGTGVTSGKRCQCPWNS-----QYPHTSHFTALRPFLNGHSHYCRNFGN 367
QY 61 RRR-PCWY 67
DB 368 QKEAPWCF 375

RESULT 43
Q15146
ID Q15146 PRELIMINARY; PRT; 810 AA.
AC Q15146
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Plasminogen precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tissue=Liver;
RA Browne M.J., Chapman C.G., Dodd I., Carey J.E., Lawrence G.M.P.,
RA Mitchell D., Robinson J.H.;
RT "Expression of recombinant human plasminogen and aglycoplasminogen in
RT HeLa cells";
RL Fibrinolysis 0:0-0(1991).
CC -!- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
DR EMBL; M74220; AAA36451.1; -.
DR HSP; P00747; 2PKA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0003809; F:thrombin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0007596; P:blood coagulation; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00051; kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 5.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 5.
DR PROSITE; PS00070; TRYPsin DOM; 1.
DR PROSITE; PS00134; TRYPsin HIS; 1.
DR PROSITE; PS00135; TRYPsin SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 810 PLASMINOGEN.

SQ SEQUENCE 810 AA; 90555 MW; B05C7D4B0D020B3C CRC64;

Query Match
Best Local Similarity 28.2%; Score 144; DB 4; Length 810;
Matches 33; Conservative 11; Mismatches 30; Indels 14; Gaps 5

QY 3 CYEGNGHFRGKASTDTMGRPCLPWNATVLTQTYHAHR---SDALQLGLGHKHYCRNPD 59
DB 377 CYHGDQSGYRGSTSTTTTKKCKQSWSS-----MTPRHQKTPENYPNAGL-TMYTCRNP 430
QY 60 NRRRPWCYVQVGLKPLV--QECWVHDC 85
DB 431 ADKGPWCFT--TDFSVRWEYCNLKKS 455

RESULT 44
Q90ZN6
ID Q90ZN6 PRELIMINARY; PRT; 709 AA.
AC Q90ZN6
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hepatocyte growth factor-like 1.
OS MST1 OR HGFL1.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Bassett D.I., Wilson S.W.;
RA "Early expression of zebrafish Hepatocyte Growth Factor-like 1
RT suggests a conserved role in vertebrate neural induction.";
RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
DR EMBL; AF370035; AAK54207.1; -.
DR HSP; P00761; IAN1.
DR ZFIN; ZDB-GENE-020806-3; mst1.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00051; kringle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; TRYPsin DOM; 1.
DR PROSITE; PS00240; TRYPsin HIS; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
SQ SEQUENCE 709 AA; 81271 MW; 9907236C5DB73A20 CRC64;

Query Match
Best Local Similarity 28.1%; Score 143.5; DB 13; Length 709;
Matches 35; Conservative 9; Mismatches 31; Indels 17; Gaps 5

QY 1 KTCVEGNHFRGKASTDTMGRPCLPWNATVLTQTYHAHR-----SDALQLGLGHKHYC 55
DB 109 KCVKGVKGYRGKVTSTTSGRTCCQWNS-----KFPDHRHTWTPSATNGLEL-----NYC 158
```



```

QY 56 RNPD-NRRRPPWCVVQVGLKPLVQECMVHDCAD 86
DB 159 RNPQDRIQFWCYT-TDPRRYESCNIPOCKD 189

RESULT 45
Q17576
ID Q17576 PRELIMINARY; PRT; 902 AA.
AC Q17576;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE KIN-8 protein (Receptor tyrosine kinase).
GN CO1G6.8 OR KIN-8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Berks M.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
RN [3]
RP SEQUENCE FROM N.A.
RA Mortimore B.J.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=N2;
RA Koga M.; Take-uchi M., Tameishi T., Ohshima Y.;
RT "Control of DAF-7 TGF expression and neuronal process development by a
RT receptor tyrosine kinase KIN-8 in C. elegans.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; Z35595; CA84639.2; -.
DR EMBL; Z47808; CA84639.2; JOINED.
DR EMBL; Z47808; CAB61003.1; -.
DR EMBL; Z35595; CAB61003.1; JOINED.
DR EMBL; AJ132947; CAC29085.1; -.
DR PIR; T18840; T18840.
DR HSSP; P11362; IFGK.
DR WormPep; CO1G6.8b; CE24774.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transmembrane receptor activity; IEA.
DR GO; GO:0004888; F:transmembrane receptor activity; IEA.
DR GO; GO:0007275; P:development; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000024; Fz domain.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig c2.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000395; Kringle; 1.
DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00219; Tyrc; 1.

QY 56 RNPD-NRRRPPWCVVQVGLKPLVQECMVHDCAD 86
DB 159 RNPQDRIQFWCYT-TDPRRYESCNIPOCKD 189

RESULT 46
Q9BLY1
ID Q9BLY1 PRELIMINARY; PRT; 928 AA.
AC Q9BLY1;
DT 01-JUN-2001 (TremBLrel. 17, Created)
DT 01-JUN-2001 (TremBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Receptor tyrosine kinase (C. elegans Cam-1 protein) (corresponding
DE sequence CO1G6.8a).
GN CO1G6.8 OR KIN-8 OR CAM-1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N2;
RA Koga M.; Take-uchi M., Tameishi T., Ohshima Y.;
RT "Control of DAF-7 TGF expression and neuronal process development by a
RT receptor tyrosine kinase KIN-8 in C. elegans.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Berks M.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
RN [4]
RP SEQUENCE FROM N.A.
RA Mortimore B.J.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; Z35595; CAC29084.1; -.
DR EMBL; AJ132946; CAC29084.1; -.
DR EMBL; Z47808; CAC29084.1; JOINED.
DR EMBL; Z47808; CAC29084.1; JOINED.
DR HSSP; P11362; IFGK.
DR WormPep; CO1G6.8a; CE32563.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0006468; P:transmembrane receptor activity; IEA.

```

DR GO; GO:0005975; P:carbohydrate metabolism; IEA.  
DR GO; GO:0007275; P:development; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR000577; FGGY\_kin.  
DR InterPro; IPR000024; Fz\_domain.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003598; Ig\_c2.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR InterPro; IPR008266; Tyr\_pkinase\_AS.  
DR Pfam; PF00047; Ig\_1.  
DR Pfam; PF00051; kringle; 1.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00018; KRINGLE.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000395; Kringle; 1.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR SMART; SM00408; IgC2; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00213; TyRK; 1.  
DR PROSITE; PS00933; FGGY\_KINASES\_1; 1.  
DR PROSITE; PS0038; FZ; 1.  
DR PROSITE; PS00835; IG LIKE; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS00070; KRINGLE\_2; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; IMMUNOGLOBULIN\_DOMAIN; Kinase; 1.  
KW ATP-binding; Immunoglobulin domain; Kinase; Kringle; Transferase;  
KW Tyrosine-protein kinase.  
SQ SEQUENCE 928 AA; 103864 MW; F13B8C9CAB30D20 CRC64;  
  
Query Match 28.0%; Score 143; DB 5; Length 928;  
Best Local Similarity 36.5%; Pred. No. 5.4e-08;  
Matches 31; Conservative 12; Mismatches 30; Indels 12; Gaps 5;  
  
QY 3 CYEGNGHYRGKASTDTMGRPCLPWNATVLOQTYHAHRSDALQGLGKHYCRNPDR 62  
Db 354 CYVNGSTQYEGVTAQTSKQCAPWIDST--SRDFNVHRFPPEL---VNSKNYCRNPGGKK 408  
  
QY 63 -RPWCYVQVGLKPLVQ--CMVHDC 84  
Db 409 SRPWCY---SKPMQGEYCDVPQC 429  
  
RESULT 47  
Q7TP84 PRELIMINARY; PRT; 759 AA.  
AC Q7TP84; (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Ab1-346.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Xu C.S., Li W.Q., Li Y.C., Yan H.M., Chang C.F., Zhao L.F., Ma H.,  
RA Wang L., Wang S.F., Han H.P., Wang G.P., Chai L.Q., Yuan J.Y.,  
RA Yang K.J., Shi J.B., Rahman S., Wang Q.N., Zhang J.B.,  
RT "Liver regeneration after PH."  
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY25159; AAP92560.1; --  
SQ SEQUENCE 759 AA; 86056 MW; 1422BFAC05C6DFA7 CRC64;  
  
Query Match 27.5%; Score 140.5; DB 11; Length 759;  
Best Local Similarity 28.7%; Pred. No. 8.6e-08;  
Matches 37; Conservative 12; Mismatches 27; Indels 53; Gaps 6;  
  
QY 3 CYEGNGHYRGKASTDTMGRPCLPWNATVLOQTYHAHRSDALQGLGKHYCRNPDR 60  
Db 28 CHNSGQSYRGTYFTVTGRTCOANSSMTPHQHSRTPEKYVNDGL-----ISNYCRNPDC 82  
  
QY 61 RRPWCYVQVGLKPLV--QECMVHDCAD 86  
Db 83 SAGPWCYT---TDPNVRYEYCNLTRCSD 107  
  
RESULT 49  
Q8AV69 PRELIMINARY; PRT; 930 AA.  
AC Q8AV69;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Receptor tyrosine kinase Xr2r2.  
GN XKR2.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Db 384 CYQNGKSYRGTSSTNTGKQSW-----VSMTPHSHSKTPANPPDAYQNLIIQWRPS 438  
QY 53 -----NYCRNPDR-RRRPWCYVQVGLKPLV--Q 77  
Db 439 VYVMSLWETFLHTHNQKYSVFLKXSSGLENNYCRNPDRDQGRPWCT---TDFSVRWE 495  
QY 78 ECVVHDCAD 86  
Db 496 YCNLKRCE 504  
  
RESULT 48  
Q16609 PRELIMINARY; PRT; 132 AA.  
AC Q16609;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE (APOARGC).  
GN APOARGC.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC Tissue=Liver;  
RA MEDLINE=95268939; PubMed=7749817;  
RA Byrne C.D., Schwartz K., Lawn R.M.;  
RT "Loss of a splice donor site at a 'skipped exon' in a gene homologous  
RT to apolipoprotein(a) leads to an mRNA encoding a protein consisting of  
RT a single kringle domain";  
RL Arterioscler. Thromb. Vasc. Biol. 15:65-70(1995).  
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  
DR EMBL; U19518; AAA85693.1; --  
DR HSSP; P00747; 1PMK.  
DR Genew; HGNC:6669; LPAL2.  
DR InterPro; IPR000001; Kringle.  
DR Pfam; PF00051; kringle; 1.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00130; KR; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS00070; KRINGLE\_2; 1.  
KW Glycoprotein; Kringle.  
SQ SEQUENCE 132 AA; 14886 MW; 3794AD30A586DBBA CRC64;  
  
Query Match 27.5%; Score 140; DB 4; Length 132;  
Best Local Similarity 38.6%; Pred. No. 1.5e-08;  
Matches 34; Conservative 7; Mismatches 35; Indels 12; Gaps 4  
  
QY 3 CYEGNGHYRGKASTDTMGRPCLPWNATVLOQTYHAHRSDALQGLGKHYCRNPDR 60  
Db 28 CHNSGQSYRGTYFTVTGRTCOANSSMTPHQHSRTPEKYVNDGL-----ISNYCRNPDC 82  
  
QY 61 RRPWCYVQVGLKPLV--QECMVHDCAD 86  
Db 83 SAGPWCYT---TDPNVRYEYCNLTRCSD 107  
  
RESULT 49  
Q8AV69 PRELIMINARY; PRT; 930 AA.  
AC Q8AV69;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Receptor tyrosine kinase Xr2r2.  
GN XKR2.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22286220; PubMed=12399314;  
 RA Hikasa H., Shibata M., Hiratani I., Taira M.;  
 RT "The Xenopus receptor tyrosine kinase Xr2 modulates morphogenetic  
 RT movements of the axial mesoderm and neuroectoderm via Wnt  
 RT signalling.";  
 RL Development 129:5227-5239(2002).  
 DR EMBL; AB087137; BAC16209.1; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005524; P:ATP binding; IEA.  
 DR GO; GO:004674; P:protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0004888; P:transmembrane receptor activity; IEA.  
 DR GO; GO:0007275; P:development; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR000024; Fz domain.  
 DR InterPro; IPR003599; Ig-like.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig\_c2.  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR000719; Prot kinase.  
 DR InterPro; IPR002290; Ser thr kinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR InterPro; IPR008266; Tyr\_pkinase\_AS.  
 DR Pfam; PF01392; Fz; 1.  
 DR Pfam; PF00047; IS; 1.  
 DR Pfam; PF00051; Kringle; 1.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00018; KRINGLE.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000395; Kringle; 1.  
 DR SMART; SM00001; Prot\_kinase; 1.  
 DR SMART; SM00409; IG; 1.  
 DR SMART; SM00408; IGc2; 1.  
 DR SMART; SM00130; KR; 1.  
 DR SMART; SM00220; S\_TK; 1.  
 DR SMART; SM00219; TyK; 1.  
 DR PROSITE; PS50038; FZ; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 DR PROSITE; PS00021; KRINGLE\_1; 1.  
 DR PROSITE; PS50070; KRINGLE\_2; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 KW Kinase.  
 SQ SEQUENCE 930 AA; 104081 MW; C68454572411A8B6 CRC64;  
 Query Match 27.5%; Score 140; DB 13; Length 930;  
 Best Local Similarity 39.4%; Pred. No. 1.2e-07;  
 Matches 28; Conservative 8; Mismatches 25; Indels 10; Gaps 4;  
 QY 3 CYEGNGHFGKASTDTMGRCLPWNATVLQQTTHAH--RSDALQLGLGKHNYCRNP 59  
 Db 312 CYNGTGYRGSVSTKSHQCQPN-----HQVSHSLSNADYFPIG-GGSHYCRNP 365  
 QY 60 NRER-PWCYVQ 69  
 Db 366 GQMEGFWCFTQ 376  
 RESULT 50  
 Q90865 PRELIMINARY; PRT; 704 AA.  
 AC Q90865;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hepatocyte growth factor-like/macrophage stimulating protein.  
 GN HGFI/NSP.  
 OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96029010; PubMed=7554499;  
 RA Thery C., Sharpe M.J., Batley S.J., Stern C.D., Gherardi E.;  
 RT "Expression of HGF/SF, HGFI/MSP and c-met suggests new functions  
 RT during early chick development.";  
 RL Dev. Genet. 17:90-101(1995)  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -1- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.  
 DR HMBL; X84043; CAAS8862.1; -.  
 DR HSSP; P00747; ICEA.  
 DR MEROPS; S01.977; -.  
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
 DR GO; GO:0008233; F:peptidase activity; IEA.  
 DR GO; GO:0004295; F:trypsin activity; IEA.  
 DR GO; GO:0008508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR003014; PAN.  
 DR InterPro; IPR003609; Pan\_app.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR Pfam; PF00051; Kringle; 4.  
 DR Pfam; PF00024; PAN; 1.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PRINTS; PR00018; KRINGLE.  
 DR ProDom; PD000395; Kringle; 4.  
 DR SMART; SM00130; KR; 4.  
 DR SMART; SM00473; PAN\_AP; 1.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS00021; KRINGLE\_1; 4.  
 DR PROSITE; PS50070; KRINGLE\_2; 4.  
 DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
 KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.  
 SQ SEQUENCE 704 AA; 79341 MW; CABD08C41367C37 CRC64;  
 Query Match 27.2%; Score 138.5; DB 13; Length 704;  
 Best Local Similarity 38.9%; Pred. No. 1.4e-07;  
 Matches 35; Conservative 9; Mismatches 33; Indels 13; Gaps 6  
 QY 1 KTCYEGNGHFGKASTDTMGRCLPWNATVLQQTTHAH-SDALQLGLGKHNYCRNP 59  
 Db 106 RECIVANGSYRGTTRTTERGLRCQHWQATP-----HDFRFLPSLNLG-ENYCRNP 159  
 QY 60 -NRERPCVYQVGLKFLV--QSCMVHDCAD 86  
 Db 160 RDKRGFWCYT---VDPNVRHQSGIKKCED 186  
 Search completed: May 25, 2004, 14:57:10  
 Job time : 17.5767 secs

APPLICANT: NIWA, MINEO  
APPLICANT: SAITO, YOSHIMASA  
APPLICANT: SASAKI, HITOSHI  
APPLICANT: HAYASHI, MASAKO  
APPLICANT: NOTANI, JOUJI  
APPLICANT: KOBAYASHI, MASAKAZU  
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: P.C.  
ADDRESS: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/811,949  
FILING DATE: 05-MAR-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 18-966-0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 61:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 354 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
J08-811-949-61

Very Match 42.2%; Score 215; DB 2; Length 354;  
Seq Local Similarity 47.0%; Pred. No. 9e-18; Mismatches 36; Indels 0; Gaps 0;  
atches 39; Conservative 8;

C 3 CYEGNGHFYRGKASDTMGRPCLPWNSATVLQQTYHAHRSDALQLGLGKHNYCENPDNRR 62  
D 8 CYFGNGSAYRGTHSLTSSGASCLPWSNMLICKVYTAQNPSAQLGLGKHNYCENPDGDA 67  
Q 63 RPWCYVQVGLKPLVQECMWHCA 85  
D 68 KPWCHVLXNRRLTWECYCDVPSCS 90

Search completed: May 25, 2004, 15:00:00  
Time: 8.0929 secs

CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESSEE: P.C.  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/811,949  
FILING DATE: 05-MAR-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 18-966-0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 63:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 472 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-811-949-63

Query Match 42.4%; Score 216; DB 2; Length 472;  
Best Local Similarity 43.7%; Pred. No. 9.6e-18;  
Matches 38; Conservative 12; Mismatches 37; Indels 0; Gaps 0;  
Q 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDR 60  
D 35 RSCYEDQGISYRGVTTTAEAGCTNNSSALAKPKYSGRPPDIRLGLGKHNYCRNPDR 94  
Q 61 RRPWCYVQVGLKPLVQECMVHDCAG 87  
D 95 DSKPWCYVFKAGYSSEFCSTACSEG 121

US-9-411-977-3  
Sequence 3, Application US/09411977  
Patent No. 6372473  
GENERAL INFORMATION:  
APPLICANT: Moore, Paul A.  
APPLICANT: Ruben, Steven M.  
APPLICANT: Ebner, Reinhard  
TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protease  
FILE REFERENCE: PF378P1  
CURRENT APPLICATION NUMBER: US/09/411,977  
CURRENT FILING DATE: 1999-10-04  
EARLIER APPLICATION NUMBER: 09/084,491  
EARLIER FILING DATE: 1998-05-27  
EARLIER APPLICATION NUMBER: 60/048,000  
EARLIER FILING DATE: 1997-05-28  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: Patent in Ver. 2.0  
SEQ ID NO 3  
LENGTH: 326  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-9-411-977-3

Query Match 42.2%; Score 215; DB 4; Length 326;  
Best Local Similarity 47.0%; Pred. No. 8.1e-18;  
Matches 39; Conservative 8; Mismatches 36; Indels 0; Gaps 0;

QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDR 62  
DB 25 CVFNGSAYRGVTHSLTESGASCLPWNMILIGKVTAQNPSAQAALGLGKHNYCRNPDR 84  
QY 63 RRPWCYVQVGLKPLVQECMVHDCAG 85  
DB 85 KPWCHVLKRRRLTWEYCDVPSCS 107  
RESULT 49  
US-08-811-949-1  
Sequence 1, Application US/08811949  
Patent No. 5840533  
GENERAL INFORMATION:  
APPLICANT: NIWA, MINEO  
APPLICANT: SAITO, YOSHIMASA  
APPLICANT: SASAKI, HITOSHI  
APPLICANT: HAYASHI, MASAKO  
APPLICANT: NOTANI, JOUJI  
APPLICANT: KOBAYASHI, MASAKAZU  
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESS: P.C.  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/811,949  
FILING DATE: 05-MAR-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 18-966-0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 347 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-811-949-1

Query Match 42.2%; Score 215; DB 2; Length 347;  
Best Local Similarity 47.0%; Pred. No. 8.7e-18;  
Matches 39; Conservative 8; Mismatches 36; Indels 0; Gaps 0;  
QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDR 62  
DB 7 CYFNGSAYRGVTHSLTESGASCLPWNMILIGKVTAQNPSAQAALGLGKHNYCRNPDR 66  
QY 63 RRPWCYVQVGLKPLVQECMVHDCAG 85  
DB 67 KPWCHVLKRRRLTWEYCDVPSCS 89  
RESULT 50  
US-08-811-949-61  
Sequence 61, Application US/08811949  
Patent No. 5840533  
GENERAL INFORMATION:

Query Match 42.2%; Score 215; DB 2; Length 347;  
Best Local Similarity 47.0%; Pred. No. 8.7e-18;  
Matches 39; Conservative 8; Mismatches 36; Indels 0; Gaps 0;  
QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDR 62  
DB 7 CYFNGSAYRGVTHSLTESGASCLPWNMILIGKVTAQNPSAQAALGLGKHNYCRNPDR 66  
QY 63 RRPWCYVQVGLKPLVQECMVHDCAG 85  
DB 67 KPWCHVLKRRRLTWEYCDVPSCS 89  
RESULT 50  
US-08-811-949-61  
Sequence 61, Application US/08811949  
Patent No. 5840533  
GENERAL INFORMATION:

MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage  
COMPUTER: Macintosh  
OPERATING SYSTEM: Macintosh  
SOFTWARE: Microsoft Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/427,640  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/689,410  
FILING DATE: 22 APRIL 1991  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 356 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
38-427-840-8

Query Match 42.7%; Score 218; DB 1; Length 356;  
Best Local Similarity 46.4%; Pred. No. 3.9e-18;  
Matches 39; Conservative 10; Mismatches 35; Indels 0; Gaps 0;

Q 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNP DNR 61  
D 8 TCYEDQGISYRGTWSTAESGAECTWNSSALAQPKYSGRPDATRLGLGKHNYCRNP DGD 67

Q 62 RRPWCYVQVGLKPLVQECMVHDCA 85  
D 68 KPWCCHLVKNRRLTWECYDVPSCS 91

R JLT 44  
U 1676-5  
Sequence 10, Application US/07/203,047  
Best Local Similarity 45.9%; Pred. No. 6.8e-18;  
Matches 39; Conservative 10; Mismatches 36; Indels 0; Gaps 0;

Q 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNP DNR 62  
D 127 CYFGSGSAYRGTHSLTESGASCLPWNMILIGKVYTAQNPSAQLGLGKHNYCRNP DGD 186

Q 63 RPPWCYVQVGLKPLVQECMVHDCA 87  
D 187 KPWCCHLVKNRRLTWECYDVPSCS 211

R JLT 45  
U 1676-5  
Sequence 10, Application US/09/553,498  
Best Local Similarity 45.9%; Pred. No. 6.8e-18;  
Matches 39; Conservative 10; Mismatches 36; Indels 0; Gaps 0;

Q 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNP DNR 62  
D 127 CYFGSGSAYRGTHSLTESGASCLPWNMILIGKVYTAQNPSAQLGLGKHNYCRNP DGD 186

Q 63 RPPWCYVQVGLKPLVQECMVHDCA 87  
D 187 KPWCCHLVKNRRLTWECYDVPSCS 211

R JLT 45  
U 1676-5  
Sequence 10, Application US/09/553,498  
Best Local Similarity 45.9%; Pred. No. 6.8e-18;  
Matches 39; Conservative 10; Mismatches 36; Indels 0; Gaps 0;

PRIOR APPLICATION NUMBER: EP99107412.1  
PRIOR FILING DATE: 1999-04-26  
NUMBER OF SEQ ID NOS: 10  
SEQ ID NO 10  
LENGTH: 378  
TYPE: PRT  
ORGANISM: E. coli  
US-09-553-498-10

Query Match 42.5%; Score 217; DB 4; Length 378;  
Best Local Similarity 47.0%; Pred. No. 5.6e-18;  
Matches 39; Conservative 8; Mismatches 36; Indels 0; Gaps 0;

QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNP DNR 62  
Db 31 CYFGSGSAYRGTHSLTESGASCLPWNMILIGKVYTAQNPSAQLGLGKHNYCRNP DGD 90

QY 63 RPPWCYVQVGLKPLVQECMVHDCA 85  
Db 91 KPWCCHLVKNRRLTWECYDVPSCS 113

RESULT 46  
US-09-618-869-10  
Sequence 10, Application US/09618869  
Patent No. 6455279  
GENERAL INFORMATION:  
APPLICANT: Ambrosius, Dorothee  
APPLICANT: Rudolph, Rainer  
APPLICANT: Schaeffner, Joerg  
APPLICANT: Schwarz, Elisabeth  
TITLE OF INVENTION: PROCESS FOR THE PRODUCTION OF NATURALLY FOLDED AND  
TITLE OF INVENTION: SECRETED PROTEINS BY CO-SECRETION OF MOLECULAR  
TITLE OF INVENTION: CHAPERONES  
FILE REFERENCE: 20381  
CURRENT APPLICATION NUMBER: US/09/618,869  
CURRENT FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: EP99114811.5  
PRIOR FILING DATE: 1999-07-29  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: Patent in ver. 2.1  
SEQ ID NO 10  
LENGTH: 378  
TYPE: PRT  
ORGANISM: Escherichia coli  
US-09-618-869-10

Query Match 42.5%; Score 217; DB 4; Length 378;  
Best Local Similarity 47.0%; Pred. No. 5.6e-18;  
Matches 39; Conservative 8; Mismatches 36; Indels 0; Gaps 0;

QY 3 CYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNP DNR 62  
Db 31 CYFGSGSAYRGTHSLTESGASCLPWNMILIGKVYTAQNPSAQLGLGKHNYCRNP DGD 90

QY 63 RPPWCYVQVGLKPLVQECMVHDCA 85  
Db 91 KPWCCHLVKNRRLTWECYDVPSCS 113

RESULT 47  
US-08-811-949-63  
Sequence 63, Application US/08811949  
Patent No. 5840533  
GENERAL INFORMATION:  
APPLICANT: NIWA, MINO  
APPLICANT: SAITO, YOSHIMASA  
APPLICANT: SASAKI, HITOSHI  
APPLICANT: HAYASHI, MASAKO  
APPLICANT: NOTANI, JOUJI  
APPLICANT: KOBAYASHI, MASAKAZU  
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
NUMBER OF SEQUENCES: 67

Query Match 43.3%; Score 221; DB 2; Length 437;  
Best Local Similarity 47.1%; Pred. No. 2.2e-18;  
Matches 40; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

Q1 3 CYEGNGHYRGKATDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHYCRNPDRR 62  
D1 2 CYEDQGISYRGTWSTAESGAECTWNSSALAKPYSGRRPDARLGLGNHNYCRNPDRS 61  
Q2 63 RWCYVQVGLKPLVQECMVHDCADG 87  
D2 62 KPCYVFKAGYSSEFCSTPACSEG 86

RI LT 41  
US-08-811-949-55  
Sequence 55, Application US/08811949  
Patent No. 5840533  
GENERAL INFORMATION:  
APPLICANT: NIWA, MINEO  
APPLICANT: SAITO, YOSHIMASA  
APPLICANT: SASAKI, HITOSHI  
APPLICANT: HAYASHI, MASAKO  
APPLICANT: NOTANI, JOUJI  
APPLICANT: KOBAYASHI, MASAKAZU  
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
ADDRESS: P.C.  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/811.949  
FILING DATE: 05-MAR-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 18-966-0  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 55:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 437 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-811-949-55

Query Match 43.3%; Score 221; DB 2; Length 437;  
Best Local Similarity 47.1%; Pred. No. 2.2e-18;  
Matches 40; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

Q1 3 CYEGNGHYRGKATDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHYCRNPDRR 62  
D1 2 CYEDQGISYRGTWSTAESGAECTWNSSALAKPYSGRRPDARLGLGNHNYCRNPDRS 61  
Q2 63 RWCYVQVGLKPLVQECMVHDCADG 87  
D2 62 KPCYVFKAGYSSEFCSTPACSEG 86

RI LT 42  
US-08-811-949-55

US-08-811-949-57  
Sequence 57, Application US/08811949  
Patent No. 5840533  
GENERAL INFORMATION:  
APPLICANT: NIWA, MINEO  
APPLICANT: SAITO, YOSHIMASA  
APPLICANT: SASAKI, HITOSHI  
APPLICANT: HAYASHI, MASAKO  
APPLICANT: NOTANI, JOUJI  
APPLICANT: KOBAYASHI, MASAKAZU  
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
ADDRESS: P.C.  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/811.949  
FILING DATE: 05-MAR-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 18-966-0  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 437 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-811-949-57

Query Match 43.3%; Score 221; DB 2; Length 437;  
Best Local Similarity 47.1%; Pred. No. 2.2e-18;  
Matches 40; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

Q1 3 CYEGNGHYRGKATDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHYCRNPDRR 62  
D1 2 CYEDQGISYRGTWSTAESGAECTWNSSALAKPYSGRRPDARLGLGNHNYCRNPDRS 61  
Q2 63 RWCYVQVGLKPLVQECMVHDCADG 87  
D2 62 KPCYVFKAGYSSEFCSTPACSEG 86

RESULT 43  
US-08-427-640-8  
Sequence 8, Application US/08427640  
Patent No. 5658788  
GENERAL INFORMATION:  
APPLICANT: Berg et al.  
TITLE OF INVENTION: Compounds and Methods for Treatment of Thromboembolic Disor  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eli Lilly and Company  
STREET: Lilly Corporate Center  
CITY: Indianapolis  
STATE: IN  
COUNTRY: U.S.A.  
ZIP: 46285  
COMPUTER READABLE FORM:





Query Match	Score	DB	Length	Indels	Gaps
8-883-795A-38	44.3%	226	562	0	0
Best Local Similarity	47.7%				
Matches	41	Conservative	9	Mismatches	36
GENERAL INFORMATION:					
APPLICANT: Delcive, Genevieve					
APPLICANT: Awang, Gregor					
TITLE OF INVENTION: Recombinant DNA Molecules and Expression					
TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator					
NUMBER OF SEQUENCES: 39					
CORRESPONDENCE ADDRESS:					
ADDRESSEE: BERSKIN & PARR					
STREET: 40 King Street West					
CITY: Toronto					
STATE: Ontario					
COUNTRY: Canada					
ZIP: M5H 3Y2					
COMPUTER READABLE FORM:					
MEDIUM TYPE: Floppy disk					
COMPUTER: IBM PC compatible					
OPERATING SYSTEM: PC-DOS/MS-DOS					
SOFTWARE: Patent in Release #1.0, Version #1.25					
CURRENT APPLICATION DATA:					
APPLICATION NUMBER: US/08/883,795A					
FILING DATE: 27-JUN-1997					
CLASSIFICATION: 435					
ATTORNEY/AGENT INFORMATION:					
NAME: Gravelle, Micheline					
REGISTRATION NUMBER: 40,261					
REFERENCE/DOCKET NUMBER: 7841-662					
TELECOMMUNICATION INFORMATION:					
TELEPHONE: (416) 364-7311					
TELEFAX: (416) 361-1398					
INFORMATION FOR SEQ ID NO: 38:					
SEQUENCE CHARACTERISTICS:					
LENGTH: 562 amino acids					
TYPE: amino acid					
STRANDEDNESS: single					
TOPOLOGY: linear					
MOLECULE TYPE: protein (tpa)					
8-883-795A-38	44.3%	226	562	0	0
Best Local Similarity	47.7%				
Matches	41	Conservative	9	Mismatches	36
GENERAL INFORMATION:					
APPLICANT: Delcive, Genevieve					
APPLICANT: Awang, Gregor					
TITLE OF INVENTION: Recombinant DNA Molecules and Expression					
TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator					
NUMBER OF SEQUENCES: 39					
CORRESPONDENCE ADDRESS:					
ADDRESSEE: BERSKIN & PARR					
STREET: 40 King Street West					
CITY: Toronto					
STATE: Ontario					
COUNTRY: Canada					
ZIP: M5H 3Y2					
COMPUTER READABLE FORM:					
MEDIUM TYPE: Floppy disk					
COMPUTER: IBM PC compatible					
OPERATING SYSTEM: PC-DOS/MS-DOS					
SOFTWARE: Patent in Release #1.0, Version #1.25					
CURRENT APPLICATION DATA:					
APPLICATION NUMBER: US/08/883,795A					
FILING DATE: 27-JUN-1997					
CLASSIFICATION: 435					
ATTORNEY/AGENT INFORMATION:					
NAME: Gravelle, Micheline					
REGISTRATION NUMBER: 40,261					
REFERENCE/DOCKET NUMBER: 7841-662					
TELECOMMUNICATION INFORMATION:					
TELEPHONE: (416) 364-7311					
TELEFAX: (416) 361-1398					
INFORMATION FOR SEQ ID NO: 38:					
SEQUENCE CHARACTERISTICS:					
LENGTH: 562 amino acids					
TYPE: amino acid					
STRANDEDNESS: single					
TOPOLOGY: linear					
MOLECULE TYPE: protein (tpa)					
8-883-795A-38	44.3%	226	562	0	0
Best Local Similarity	47.7%				
Matches	41	Conservative	9	Mismatches	36
GENERAL INFORMATION:					
APPLICANT: Delcive, Genevieve					
APPLICANT: Awang, Gregor					
TITLE OF INVENTION: Recombinant DNA					

[illegible]



```

; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; 18-811-949-39
;
; Query Match 44.3%; Score 226; DB 2; Length 527;
; st Local Similarity 47.7%; Pred. No. 6.9e-19;
; tches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;
;
; Q 2 TCYEGNGHYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRPNDR 61
; D 91 TCYEDQGISYRGTSWTAESGAECTWNSSALQAQKPYSGRRPDALRLGLGNHNYCRPNDR 150
;
; Q 62 RRPWCYVQVGLKPLVQECWVHDCADG 87
; D 151 SKPCYVFKAGYSSEFCSTPACSEG 176
;
;
; LT 28
; PC US91-01025A-2
; Sequence 2, Application PC/TUS9101025A
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: Tissue Plasminogen Activator Having Fibrin
; TITLE OF INVENTION: Specific Properties
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/01025A
; FILING DATE: 19910214
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/486,657
; FILING DATE: 1 March 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 454P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/266-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; US91-01025A-2

```

```

; Query Match 44.3%; Score 226; DB 5; Length 527;
; st Local Similarity 47.7%; Pred. No. 6.9e-19;
; tches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

```

```

; QY 2 TCYEGNGHYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRPNDR 61
; DB 91 TCYEDQGISYRGTSWTAESGAECTWNSSALQAQKPYSGRRPDALRLGLGNHNYCRPNDR 150
;
; QY 62 RRPWCYVQVGLKPLVQECWVHDCADG 87
; DB 151 SKPCYVFKAGYSSEFCSTPACSEG 176
;
; RESULT 29
; 5185259-8
; ; APPLICANT: GOEDDEL, DAVID V.; KOHR, WILLIAM J.; PENNICA, DIANE;
; ; VEHAR, GORDON A.
; ; TITLE OF INVENTION: TRUNCATED HUMAN TISSUE PLASMINOGEN
; ; ACTIVATOR
; ; NUMBER OF SEQUENCES: 15
; ; CURRENT APPLICATION DATA:
; ; APPLICATION NUMBER: US/07/489,855
; ; FILING DATE: 02-MAR-1990
; ; PRIOR APPLICATION DATA:
; ; APPLICATION NUMBER: 12,694
; ; FILING DATE: 09-FEB-1987
; ; APPLICATION NUMBER: 483,052
; ; FILING DATE: 07-APR-1983
; ; APPLICATION NUMBER: 398,003
; ; FILING DATE: 14-JUL-1982
; ; APPLICATION NUMBER: 374,860
; ; FILING DATE: 05-MAY-1982
; ; SEQ ID NO: 8:
; ; LENGTH: 527
; 5185259-8
;
; Query Match 44.3%; Score 226; DB 6; Length 527;
; Best Local Similarity 47.7%; Pred. No. 6.9e-19;
; Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;
;
; QY 2 TCYEGNGHYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRPNDR 61
; DB 91 TCYEDQGISYRGTSWTAESGAECTWNSSALQAQKPYSGRRPDALRLGLGNHNYCRPNDR 150
;
; QY 62 RRPWCYVQVGLKPLVQECWVHDCADG 87
; DB 151 SKPCYVFKAGYSSEFCSTPACSEG 176
;
; RESULT 30
; 5520913-1
; ; APPLICANT: ANDERSON, STEPHEN; BENNETT, WILLIAM F.; BOTSTEIN,
; ; DAVID; HIGGINS, DEBORAH L.; PRONTI, NICHOLAS F.; ZOLLER, MARK J.
; ; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR HAVING
; ; ZYMOGENIC PROPERTIES
; ; NUMBER OF SEQUENCES: 35
; ; CURRENT APPLICATION DATA:
; ; APPLICATION NUMBER: US/08/88,451
; ; FILING DATE: 06-JUL-1993
; ; PRIOR APPLICATION DATA:
; ; APPLICATION NUMBER: 770,510
; ; FILING DATE: 03-OCT-1991
; ; APPLICATION NUMBER: 384,608
; ; FILING DATE: 24-JUL-1989
; ; APPLICATION NUMBER: 240,856
; ; FILING DATE: 02-SEP-1988
; ; SEQ ID NO: 1:
; ; LENGTH: 527
; 5520913-1
;
; Query Match 44.3%; Score 226; DB 6; Length 527;
; Best Local Similarity 47.7%; Pred. No. 6.9e-19;
; Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;
;
; QY 2 TCYEGNGHYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRPNDR 61

```

```

atches 84; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Q 1 KTCYEGNGHFYRGKASTDTMGSPCLPWSATVLOQTYHAHRSALQGLGKHNKCRPN 60
D 48 KTCYEGNGHFYRGKASTDTMGSPCLPWSATVLOQTYHAHRSALQGLGKHNKCRPN 107
Q 61 RRPWCYVQVGLKPLVQECMVHDCADGK 88
D 108 RRPWCYVQVGLKPLVQECMVHDCADGK 135

R JLT 25
U 08-560-098A-51
; Sequence 51, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WENDET, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; NUMBER OF SEQUENCES: 60
; COAGULATION-INHIBITING PROPERTIES
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evensen, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560.098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 477 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
U 08-560-098A-51

ery Match 47.3%; Score 241; DB 2; Length 477;
st Local Similarity 50.0%; Pred. No. 9.8e-21;
atches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0;
Q 2 TCYEGNGHFYRGKASTDTMGSPCLPWSATVLOQTYHAHRSALQGLGKHNKCRPN 61
D 127 TCYEGNGHFYRGKASTDTMGSPCLPWSATVLOQTYHAHRSALQGLGKHNKCRPN 186
Q 62 RRPWCYVQVGLKPLVQECMVHDCADGK 85
D 187 RRPWCYVQVGLKPLVQECMVHDCADGK 210

R JLT 26
U 07-609-510B-16
; Sequence 16, Application US/07609510B
; Patent No. 5326700

```

```

; GENERAL INFORMATION:
; APPLICANT: Berg et al.
; TITLE OF INVENTION: Method for Altering Post-Translational Processing of Tissue P;
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN.
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/609.510B
; FILING DATE: 19901106
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-609-510B-16

Query Match 44.3%; Score 226; DB 1; Length 527;
Best Local Similarity 47.7%; Pred. No. 6.9e-19;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;
QY 2 TCYEGNGHFYRGKASTDTMGSPCLPWSATVLOQTYHAHRSALQGLGKHNKCRPN 61
Db 91 TCYEQGHSYRGKASTDTMGSPCLPWSATVLOQTYHAHRSALQGLGKHNKCRPN 150
QY 62 RRPWCYVQVGLKPLVQECMVHDCADGK 87
Db 151 RRPWCYVQVGLKPLVQECMVHDCADGK 176

RESULT 27
US-08-811-949-39
; Sequence 39, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811.949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

```

48 KTCYEGNGHFGKASDTDTGRCLPWN SATVLQOYTHAHRSDALQGLGKHNYCRNPDN 107

61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88

108 RRRPWCYVQVGLKLLVQECMVHDCADGK 135

REF LT 22

US 18-560-098A-48

Sequence 48, Application US/08560098A

Patent No. 5976841

GENERAL INFORMATION:

APPLICANT: WENNDT, Stephan

APPLICANT: HEINZEL-WIELAND, Regina

APPLICANT: STEFFENS, Gerd Josef

TITLE OF INVENTION: Proteins having Fibrinolytic and

TITLE OF INVENTION: Coagulation-inhibiting Properties

NUMBER OF SEQUENCES: 60

CORRESPONDENCE ADDRESS:

ADDRESSEE: Evenson, McKeown, Edwards & Lenahan

STREET: 1200 G Street, N.W., Suite 700

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/560,098A

FILING DATE: 17-NOV-1995

PRIOR APPLICATION NUMBER: P 44 40 892.7

FILING DATE: 17-NOV-1994

ATTORNEY/AGENT INFORMATION:

NAME: EVANS, Joseph D.

REGISTRATION NUMBER: 26,269

REFERENCE/DOCKET NUMBER: 148/42448

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 628-8800

TELEFAX: (202) 628-8844

INFORMATION FOR SEQ ID NO: 48:

SEQUENCE CHARACTERISTICS:

LENGTH: 411 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US 18-560-098A-48

ery Match 97.8%; Score 499; DB 2; Length 411;

st Local Similarity 98.9%; Pred. No. 9.6e-52;

tches 87; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY

1 KTCYEGNGHFGKASDTDTGRCLPWN SATVLQOYTHAHRSDALQGLGKHNYCRNPDN 60

48 KTCYEGNGHFGKASDTDTGRCLPWN SATVLQOYTHAHRSDALQGLGKHNYCRNPDN 107

DY

61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88

108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

REF LT 23

US 9-101-272G-62

Sequence 62, Application US/09101272G

Patent No. 6509445

GENERAL INFORMATION:

APPLICANT: Nissin Food Products Co., Ltd.

TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR

FILE REFERENCE: Q50979

CURRENT APPLICATION NUMBER: US/09/101,272G

CURRENT FILING DATE: 1998-07-08

PRIOR APPLICATION NUMBER: JP 1059/1996

PRIOR FILING DATE: 1996-01-08

NUMBER OF SEQ ID NOS: 107

SOFTWARE: Patentin version 3.1

SEQ ID NO 62

LENGTH: 89

TYPE: PRI

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: residues 43-131 of the ATF domain of uPA

US-09-101-272G-62

Query Match 95.9%; Score 489; DB 4; Length 89;

Best Local Similarity 100.0%; Pred. No. 2.5e-51;

Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 KTCYEGNGHFGKASDTDTGRCLPWN SATVLQOYTHAHRSDALQGLGKHNYCRNPDN 60

6 KTCYEGNGHFGKASDTDTGRCLPWN SATVLQOYTHAHRSDALQGLGKHNYCRNPDN 65

DY

61 RRRPWCYVQVGLKPLVQECMVHDC 84

66 RRRPWCYVQVGLKPLVQECMVHDC 89

RESULT 24

US-08-142-590B-25

Sequence 25, Application US/08142590B

Patent No. 6120765

GENERAL INFORMATION:

APPLICANT: HIBINO, Tashihiko; TAKAHASHI, Tadahito; HORII, Izumi; and

TITLE OF INVENTION: UROKINASE PLASMINOGEN ACTIVATOR FRAGMENTS

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD, LLP

STREET: 28 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/142,590B

FILING DATE: 25-OCT-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/042,318

FILING DATE: 02-APR-1993

ATTORNEY/AGENT INFORMATION:

NAME: Myers, Paul L.

REGISTRATION NUMBER: 35,965

REFERENCE/DOCKET NUMBER: MGP-009CP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 157 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: internal

US-08-142-590B-25

Query Match 95.3%; Score 486; DB 3; Length 157;

Best Local Similarity 95.5%; Pred. No. 1.1e-50;

APPLICANT: STEPPENS, Gerd Josef  
TITLE OF INVENTION: Proteins having fibrinolytic and  
TITLE OF INVENTION: Coagulation-inhibiting Properties  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan  
STREET: 1200 G Street, N.W., Suite 700  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30 (BPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/560,098A  
FILING DATE: 17-NOV-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: P 44 40 892.7  
FILING DATE: 17-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: EVANS, Joseph D.  
REGISTRATION NUMBER: 26,269  
REFERENCE/DOCKET NUMBER: 148/42448  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 628-8800  
TELEFAX: (202) 628-8844  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 432 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US 08-560-098A-47

Query Match 100.0%; Score 510; DB 2; Length 432;  
Best Local Similarity 100.0%; Pred. No. 4.9e-53;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 60  
69 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 128  
61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
129 RRRPWCYVQVGLKPLVQECMVHDCADGK 156

R 1LT 19  
U 09-101-272G-80  
Sequence 80, Application US/09101272G  
Patent No. 6509445  
GENERAL INFORMATION:  
APPLICANT: Nissin Food Products Co., Ltd.  
TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR  
FILE REFERENCE: Q50979  
CURRENT APPLICATION NUMBER: US/09/101,272G  
CURRENT FILING DATE: 1998-07-08  
PRIOR APPLICATION NUMBER: JP 1059/1996  
PRIOR FILING DATE: 1996-01-08  
NUMBER OF SEQ ID NOS: 107  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 80  
LENGTH: 194  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: ATPHI chimeric protein  
U 09-101-272G-80

Query Match 99.0%; Score 505; DB 4; Length 194;  
Best Local Similarity 100.0%; Pred. No. 7.5e-53;  
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 60  
DB 49 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 108  
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADG 87  
DB 109 RRRPWCYVQVGLKPLVQECMVHDCADG 135

RESULT 20  
US-09-101-272G-96  
Sequence 96, Application US/09101272G  
Patent No. 6509445  
GENERAL INFORMATION:  
APPLICANT: Nissin Food Products Co., Ltd.  
TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR  
FILE REFERENCE: Q50979  
CURRENT APPLICATION NUMBER: US/09/101,272G  
CURRENT FILING DATE: 1998-07-08  
PRIOR APPLICATION NUMBER: JP 1059/1996  
PRIOR FILING DATE: 1996-01-08  
NUMBER OF SEQ ID NOS: 107  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 96  
LENGTH: 201  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: ATPHI-CL chimeric protein  
US-09-101-272G-96

Query Match 99.0%; Score 505; DB 4; Length 201;  
Best Local Similarity 100.0%; Pred. No. 7.9e-53;  
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 60  
DB 49 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 108  
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADG 87  
DB 109 RRRPWCYVQVGLKPLVQECMVHDCADG 135

RESULT 21  
US-09-181-816-1  
Sequence 1, Application US/09181816  
Patent No. 6277818  
GENERAL INFORMATION:  
APPLICANT: MAZAR, Andrew P.  
TITLE OF INVENTION: CYCLIC PEPTIDE LIGANDS THAT TARGET UROKINASE  
FILE REFERENCE: 329042000300 SIDN 1-7  
CURRENT APPLICATION NUMBER: US/09/181,816  
CURRENT FILING DATE: 1998-10-29  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 411  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-181-816-1

Query Match 98.0%; Score 500; DB 3; Length 411;  
Best Local Similarity 98.9%; Pred. No. 7.3e-52;  
Matches 87; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 60



Q: 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDN 60  
D: 48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDN 107  
Q: 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
D: 108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

RI LT 12  
U: 18-153-799-18  
; Sequence 18, Application US/08153799  
; Patent No. 576883  
; GENERAL INFORMATION:  
; APPLICANT: Ballance, David J  
; APPLICANT: Goodey, Andrew R  
; TITLE OF INVENTION: Polypeptides  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: R Hain Swope, BOC Health Care Inc  
; STREET: 100 Mountain Avenue  
; CITY: Murray Hill  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07974  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/153,799  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/847975  
; FILING DATE: 06-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 8909916.2  
; FILING DATE: 29-APR-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB90/00650  
; FILING DATE: 26-APR-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/775952  
; FILING DATE: 29-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Swope, R Hain  
; REGISTRATION NUMBER: 24864  
; REFERENCE/DOCKET NUMBER: 92H832  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908) 665 2400  
; TELEFAX: (908) 771 6159  
; TELEX: 219484  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 411 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
U: 18-153-799-18

Query Match 100.0%; Score 510; DB 1; Length 411;  
Best Local Similarity 100.0%; Pred. No. 4.6e-53;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Q: 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDN 60  
D: 48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDN 107  
Q: 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88

Db 108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

RESULT 13  
US-09-403-736-2  
; Sequence 2, Application US/09403736  
; Patent No. 6638502  
; GENERAL INFORMATION:  
; APPLICANT: Aventis S.A.  
; APPLICANT: Li, Hong  
; APPLICANT: LU, He  
; APPLICANT: GRISCELLI, Frank  
; APPLICANT: OPOLON, Paule  
; APPLICANT: SORIA, Claudine  
; APPLICANT: RAGOT, Thierry  
; APPLICANT: LEGRAND, Yves  
; APPLICANT: SORIA, Jeanette  
; APPLICANT: MABILAT, Christelle  
; APPLICANT: PERRICAUDET, Michel  
; APPLICANT: YEH, Patrice  
; TITLE OF INVENTION: Adenovirus-Mediated Intratumoral Delivery Of An Angiogenesis Anti  
; FILE REFERENCE: A2778A-US  
; CURRENT APPLICATION NUMBER: US/09/403,736  
; CURRENT FILING DATE: 1999-10-26  
; PRIOR APPLICATION NUMBER: PCT/EP98/02491  
; PRIOR FILING DATE: 1998-11-05  
; PRIOR APPLICATION NUMBER: 60/044,980  
; PRIOR FILING DATE: 1997-04-28  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 2  
; LENGTH: 411  
; TYPE: PRT  
; ORGANISM: humanurokinase  
US-09-403-736-2

Query Match 100.0%; Score 510; DB 4; Length 411;  
Best Local Similarity 100.0%; Pred. No. 4.6e-53;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDN 60  
Db 48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDN 107  
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
Db 108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

RESULT 14  
US-07-942-157A-3  
; Sequence 3, Application US/07942157A  
; Patent No. 5648253  
; GENERAL INFORMATION:  
; APPLICANT: Wei, Cha-Mer  
; TITLE OF INVENTION: Inhibitor-Resistant Urokinase  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Kilpatrick & Cody  
; STREET: 1100 Peachtree Street Suite 2800  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30309-4530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/942,157A  
; FILING DATE: 19920908



COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/967,024C  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: P 44 42 665.8  
FILING DATE: 30-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: EVANS, Joseph D.  
REGISTRATION NUMBER: 26,269  
REFERENCE/DOCKET NUMBER: 148/42444  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 628-8800  
TELEFAX: (202) 628-8844  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 393 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
MOLECULE TYPE: protein  
TOPOLOGY: linear  
M-0967-024C-25

Query Match 100.0%; Score 510; DB 3; Length 393;  
Best Local Similarity 100.0%; Pred. No. 4.4e-53;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQGLGKHNCRNPND 60  
3 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQGLGKHNCRNPND 62  
61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
63 RRRPWCYVQVGLKPLVQECMVHDCADGK 90

LT 10  
US-087-163-1  
Sequence 1, Application US/08087163  
GENERAL INFORMATION:  
APPLICANT: Liu, Jian-Ning  
TITLE OF INVENTION: PRO-UKINASE MUTANTS  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/087,163  
FILING DATE: 07/02/93  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Fasse, J. Peter  
REGISTRATION NUMBER: 32,983  
REFERENCE/DOCKET NUMBER: 04353/003001

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/967,024C  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: P 44 42 665.8  
FILING DATE: 30-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: EVANS, Joseph D.  
REGISTRATION NUMBER: 26,269  
REFERENCE/DOCKET NUMBER: 148/42444  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 628-8800  
TELEFAX: (202) 628-8844  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 393 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
MOLECULE TYPE: protein  
TOPOLOGY: linear  
M-0967-024C-25

Query Match 100.0%; Score 510; DB 3; Length 393;  
Best Local Similarity 100.0%; Pred. No. 4.4e-53;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQGLGKHNCRNPND 60  
3 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQGLGKHNCRNPND 62  
61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
63 RRRPWCYVQVGLKPLVQECMVHDCADGK 90

LT 10  
US-087-163-1  
Sequence 1, Application US/08087163  
GENERAL INFORMATION:  
APPLICANT: Liu, Jian-Ning  
TITLE OF INVENTION: PRO-UKINASE MUTANTS  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/087,163  
FILING DATE: 07/02/93  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Fasse, J. Peter  
REGISTRATION NUMBER: 32,983  
REFERENCE/DOCKET NUMBER: 04353/003001

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 411  
TYPE: amino acid  
STRANDEDNESS: N/A  
TOPOLOGY: N/A  
US-08-087-163-1

Query Match 100.0%; Score 510; DB 1; Length 411;  
Best Local Similarity 100.0%; Pred. No. 4.6e-53;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQGLGKHNCRNPND 60  
DB 48 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQGLGKHNCRNPND 107  
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
DB 108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

RESULT 11  
US-08-286-748B-18  
Sequence 18, Application US/08286748B  
Patent No. 5759542  
GENERAL INFORMATION:  
APPLICANT: Victor Gurewich  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY  
OF DRUGS BY PLATELETS FOR THE TREATMENT OF  
CARDIOVASCULAR AND OTHER DISEASES  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/286,748B  
FILING DATE: August 5, 1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: J. Peter Fasse  
REGISTRATION NUMBER: 32,983  
REFERENCE/DOCKET NUMBER: 04547/013001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 411  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-286-748B-18

Query Match 100.0%; Score 510; DB 1; Length 411;  
Best Local Similarity 100.0%; Pred. No. 4.6e-53;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 100.0%; Score 510; DB 1; Length 365;  
Best Local Similarity 100.0%; Pred. No. 4e-53;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q: 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60  
D: 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 61  
Q: 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
D: 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89

LT 7  
US 8-560-098A-44  
Sequence 44, Application US/08560098A  
Patent No. 5976841  
GENERAL INFORMATION:  
APPLICANT: WNEPDT, Stephan  
APPLICANT: HEINZEL-WIELAND, Regina  
APPLICANT: STEFFENS, Gerd Josef  
TITLE OF INVENTION: Proteins having Fibrinolytic and  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan  
STREET: 1200 G Street, N.W., Suite 700  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/560.098A  
FILING DATE: 17-NOV-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: P 44 40 892.7  
FILING DATE: 17-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: EVANS, Joseph D.  
REGISTRATION NUMBER: 26,269  
REFERENCE/DOCKET NUMBER: 148/42448  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 628-8800  
TELEFAX: (202) 628-8844  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 393 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-560-098A-44

Query Match 100.0%; Score 510; DB 2; Length 393;  
Best Local Similarity 100.0%; Pred. No. 4.4e-53;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q: 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60  
D: 3 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 62  
Q: 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
D: 63 RRRPWCYVQVGLKPLVQECMVHDCADGK 90

LT 8  
US 8-967-024C-24

Sequence 24, Application US/08967024C  
Patent No. 6133011  
GENERAL INFORMATION:  
APPLICANT: WNEPDT, Stephan  
APPLICANT: STEFFENS, Gerd Josef  
APPLICANT: JANOCHA, Elke  
APPLICANT: HEINZEL-WIELAND, Regina  
TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan  
STREET: 1200 G Street, N.W., Suite 700  
CITY: Washington  
STATE: DC  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/967.024C  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: P 44 42 665.8  
FILING DATE: 30-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: EVANS, Joseph D.  
REGISTRATION NUMBER: 26,269  
REFERENCE/DOCKET NUMBER: 148/42444  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 628-8800  
TELEFAX: (202) 628-8844  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 393 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-967-024C-24

Query Match 100.0%; Score 510; DB 3; Length 393;  
Best Local Similarity 100.0%; Pred. No. 4.4e-53;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60  
Db 3 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 62  
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
Db 63 RRRPWCYVQVGLKPLVQECMVHDCADGK 90

RESULT 9  
US-08-967-024C-25  
Sequence 25, Application US/08967024C  
Patent No. 6133011  
GENERAL INFORMATION:  
APPLICANT: WNEPDT, Stephan  
APPLICANT: STEFFENS, Gerd Josef  
APPLICANT: JANOCHA, Elke  
APPLICANT: HEINZEL-WIELAND, Regina  
TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan  
STREET: 1200 G Street, N.W., Suite 700  
CITY: Washington  
STATE: DC  
ZIP: 20005

128 RRRPWCYVQVGLKPLVQECMVHDCADGK 155

LT 4  
9-101-272G-98  
quence 98, Application US/09101272G  
tent No. 6509445  
NERAL INFORMATION:  
PPLICANT: Nissin Food Products Co., Ltd.  
TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR  
ILE REFERENCE: Q50979  
URRENT APPLICATION NUMBER: US/09/101.272G  
URRENT FILING DATE: 1998-07-08  
RIOR APPLICATION NUMBER: JP 1059/1996  
RIOR FILING DATE: 1996-01-08  
UMBER OF SEQ ID NOS: 107  
FTWARE: Patentin version 3.1  
Q ID NO 98  
LENGTH: 208  
YPE: PRT  
ORGANISM: Artificial Sequence  
EATURE:  
OTHER INFORMATION: ATFHI-ML chimeric protein  
9-101-272G-98

ery Match 100.0%; Score 510; DB 4; Length 208;  
st Local Similarity 100.0%; Pred. No. 2.1e-53;  
tches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPDN 60  
49 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPDN 108  
61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
109 RRRPWCYVQVGLKPLVQECMVHDCADGK 136

LT 5  
9-093-741-83  
quence 83, Application US/08093741  
tent No. 5681721  
ENERAL INFORMATION:  
APPLICANT: STEFFENS, GERD J.  
APPLICANT: WENDT, STEPHAN  
APPLICANT: SCHNEIDER, JOHANNES  
APPLICANT: HEINZEL-WIELAND, REGINA  
APPLICANT: SAUNDERS, DEREK J.  
TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH  
TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN  
NUMBER OF SEQUENCES: 83  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan  
STREET: 1200 G Street, N. W. Suite 700  
CITY: Washington, D.C.  
COUNTRY: U.S.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/093,741  
FILING DATE: 20-JUL-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE P43 23 754.1  
FILING DATE: 15-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: EVANS, Joseph D.

REGISTRATION NUMBER: 26,269  
REFERENCE/DOCKET NUMBER: 148/41345  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)628-8800  
TELEFAX: (202)628-8844  
INFORMATION FOR SEQ ID NO: 83:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 365 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-093-741-83

Query Match 100.0%; Score 510; DB 1; Length 365;  
Best Local Similarity 100.0%; Pred. No. 4e-53;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPDN 60  
Db 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRNPDN 61  
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
Db 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89

RESULT 6  
US-08-720-012-83  
Sequence 83, Application US/08720012  
Patent No. 5747251  
GENERAL INFORMATION:  
APPLICANT: STEFFENS, GERD J.  
APPLICANT: WENDT, STEPHAN  
APPLICANT: SCHNEIDER, JOHANNES  
APPLICANT: HEINZEL-WIELAND, REGINA  
APPLICANT: SAUNDERS, DEREK J.  
TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH  
TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN  
NUMBER OF SEQUENCES: 83  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan  
STREET: 1200 G Street, N. W. Suite 700  
CITY: Washington, D.C.  
COUNTRY: U.S.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/720,012  
FILING DATE: 27-SEP-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/093,741  
FILING DATE: 20-JUL-1993  
APPLICATION NUMBER: DE P43 23 754.1  
FILING DATE: 15-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: EVANS, Joseph D.  
REGISTRATION NUMBER: 26,269  
REFERENCE/DOCKET NUMBER: 148/41345  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)628-8800  
TELEFAX: (202)628-8844  
INFORMATION FOR SEQ ID NO: 83:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 365 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-720-012-83

ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: System 7.1  
SOFTWARE: Word 5.1 (PatentIn)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/797,689  
FILING DATE: 31-JAN-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/256,927  
FILING DATE: 28-JUL-1994  
FILING DATE: 31-JAN-1992  
APPLICATION NUMBER: FR 92/01064  
FILING DATE: 31-JAN-1992  
APPLICATION NUMBER: PCT/FR93/00085  
FILING DATE: 28-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith Ph.D., Julie K.  
REGISTRATION NUMBER: P-38,619  
REFERENCE/DOCKET NUMBER: S792006-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 138 amino acids  
TYPE: amino acid  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-08-797-689-12

Query Match 100.0%; Score 510; DB 2; Length 138;  
Best Local Similarity 100.0%; Pred. No. 1.3e-53;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60  
D 51 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 110

Q 61 RRRPCYVQVGLKPLVQECMVHDCADGK 89  
D 111 RRRPCYVQVGLKPLVQECMVHDCADGK 138

U 08-797-689-12

U 09-984-186-12

Query Match 100.0%; Score 510; DB 4; Length 138;  
Best Local Similarity 100.0%; Pred. No. 1.3e-53;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60  
D 51 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 110

Q 61 RRRPCYVQVGLKPLVQECMVHDCADGK 89  
D 111 RRRPCYVQVGLKPLVQECMVHDCADGK 138

U 09-984-186-12

Query Match 100.0%; Score 510; DB 4; Length 200;  
Best Local Similarity 100.0%; Pred. No. 2e-53;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60  
D 68 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 127  
Q 61 RRRPCYVQVGLKPLVQECMVHDCADGK 88

APPLICATION NUMBER: US/09/984,186  
FILING DATE: 29-Oct-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/797,689  
FILING DATE: 31-JAN-1997  
APPLICATION NUMBER: US 08/256,927  
FILING DATE: 28-JUL-1994  
APPLICATION NUMBER: FR 92/01064  
FILING DATE: 31-JAN-1992  
APPLICATION NUMBER: PCT/FR93/00085  
FILING DATE: 28-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith Ph.D., Julie K.  
REGISTRATION NUMBER: P-38,619  
REFERENCE/DOCKET NUMBER: S792006-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 138 amino acids  
TYPE: amino acid  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-09-984-186-12

Query Match 100.0%; Score 510; DB 4; Length 138;  
Best Local Similarity 100.0%; Pred. No. 1.3e-53;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60  
D 51 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 110

Q 61 RRRPCYVQVGLKPLVQECMVHDCADGK 88  
D 111 RRRPCYVQVGLKPLVQECMVHDCADGK 138

RESULT 3  
US-09-101-272G-73  
Sequence 73, Application US/09101272G  
Patent No. 6509445  
GENERAL INFORMATION:  
APPLICANT: Nissin Food Products Co., Ltd.  
TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR  
FILE REFERENCE: Q50979  
CURRENT APPLICATION NUMBER: US/09/101,272G  
CURRENT FILING DATE: 1998-07-08  
PRIOR APPLICATION NUMBER: JP 1059/1996  
PRIOR FILING DATE: 1996-01-08  
NUMBER OF SEQ ID NOS: 107  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 73  
LENGTH: 200  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: ATP domain of uPA  
US-09-101-272G-73

Query Match 100.0%; Score 510; DB 4; Length 200;  
Best Local Similarity 100.0%; Pred. No. 2e-53;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60  
D 68 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 127  
Q 61 RRRPCYVQVGLKPLVQECMVHDCADGK 88

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

rotein - protein search, using sw model

on: May 25, 2004, 14:48:05 ; Search time 6.0929 seconds  
(without alignments)  
745.636 Million cell updates/sec

US-09-880-503-1

ect score: 510

ence: 1:KTCYENGHRYGKASTDTM.....QVGLKPLVQECWVHDCADGK 88

ing table: BLOSUM62

Gapop 10.0 , Gapext 0.5

ched: 389414 seqs, 51625971 residues

l number of hits satisfying chosen parameters: 389414

mum DB seq length: 0

mum DB seq length: 2000000000

-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

base :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCRUS COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Score	Query Match	Length	ID	Description
1	510	100.0	138	2	US-08-797-689-12
2	510	100.0	138	4	US-09-984-186-12
3	510	100.0	200	4	US-09-101-272G-73
4	510	100.0	208	4	US-09-101-272G-98
5	510	100.0	365	1	US-08-093-741-83
6	510	100.0	365	1	US-08-720-012-83
7	510	100.0	393	2	US-08-560-098A-44
8	510	100.0	393	3	US-08-967-024C-24
9	510	100.0	393	3	US-08-967-024C-25
10	510	100.0	411	1	US-08-087-163-1
11	510	100.0	411	1	US-08-286-748B-18
12	510	100.0	411	1	US-08-153-799-18
13	510	100.0	411	4	US-09-403-736-2
14	510	100.0	430	1	US-07-942-157A-3
15	510	100.0	430	6	5219569-2
16	510	100.0	431	4	US-09-101-272G-1
17	510	100.0	431	6	518829-1
18	510	100.0	432	2	US-08-560-098A-47
19	505	99.0	194	4	US-09-101-272G-80
20	505	99.0	201	4	US-09-101-272G-96
21	500	98.0	411	3	US-09-181-816-1
22	499	97.8	411	2	US-08-560-098A-48
23	489	95.9	89	4	US-09-101-272G-62
24	486	95.3	157	3	US-08-442-590B-25
25	241	47.3	477	2	US-08-560-098A-51
26	226	44.3	527	1	US-07-609-510B-16
27	226	44.3	527	2	US-08-811-949-39

28	226	44.3	527	5	PCT-US91-01025A-2	Sequence 2, Appli
29	226	44.3	527	6	5185259-8	Patent No. 5185259
30	226	44.3	527	6	520313-1	Patent No. 520313
31	226	44.3	546	6	5200340-6	Patent No. 5200340
32	226	44.3	562	2	US-08-811-949-43	Sequence 43, Appl
33	226	44.3	562	2	US-08-560-098A-50	Sequence 50, Appl
34	226	44.3	562	2	US-08-883-795A-38	Sequence 38, Appl
35	226	44.3	562	4	US-09-703-695A-4	Sequence 4, Appli
36	226	44.3	562	6	5185259-3	Patent No. 5185259
37	226	44.3	562	6	5200340-2	Patent No. 5200340
38	226	44.3	562	6	5344773-2	Patent No. 5344773
39	221	43.3	437	2	US-08-811-949-49	Sequence 49, Appl
40	221	43.3	437	2	US-08-811-949-51	Sequence 51, Appl
41	221	43.3	437	2	US-08-811-949-55	Sequence 55, Appl
42	221	43.3	437	2	US-08-811-949-57	Sequence 57, Appl
43	218	42.7	356	1	US-08-427-640-8	Sequence 8, Appli
44	218	42.7	562	6	5244676-5	Patent No. 5244676
45	217	42.5	378	4	US-09-553-488-10	Sequence 10, Appl
46	217	42.5	378	4	US-09-618-869-10	Sequence 10, Appl
47	216	42.4	472	2	US-08-811-949-63	Sequence 63, Appl
48	215	42.2	326	4	US-09-411-977-3	Sequence 3, Appli
49	215	42.2	347	2	US-08-811-949-1	Sequence 1, Appli
50	215	42.2	354	2	US-08-811-949-61	Sequence 61, Appl
51	215	42.2	355	1	US-08-137-116-1	Sequence 1, Appli
52	215	42.2	355	1	US-08-217-618-1	Sequence 1, Appli
53	215	42.2	355	1	US-08-427-640-2	Sequence 2, Appli
54	215	42.2	355	1	US-08-217-617A-1	Sequence 1, Appli
55	215	42.2	355	1	US-08-217-616-1	Sequence 1, Appli
56	215	42.2	355	2	US-08-811-949-45	Sequence 45, Appl
57	215	42.2	355	2	US-08-811-949-47	Sequence 47, Appl
58	215	42.2	355	2	US-08-811-949-53	Sequence 53, Appl
59	215	42.2	355	2	US-08-811-949-59	Sequence 59, Appl
60	215	42.2	355	3	US-08-794-528-1	Sequence 1, Appli
61	215	42.2	355	6	5223256-1	Patent No. 5223256
62	215	42.2	389	2	US-08-811-949-65	Sequence 65, Appl
63	215	42.2	389	2	US-08-811-949-67	Sequence 67, Appl
64	213	41.8	83	2	US-08-811-949-2	Sequence 2, Appli
65	210	41.2	355	1	US-08-427-640-6	Sequence 6, Appli
66	200.5	39.3	356	1	US-08-427-640-4	Sequence 4, Appli
67	193	37.8	655	1	US-08-148-910-12	Sequence 12, Appl
68	193	37.8	655	1	US-08-448-937A-12	Sequence 12, Appl
69	154	30.2	458	3	US-09-039-609-2	Sequence 2, Appli
70	149	29.2	812	4	US-08-991-761A-7	Sequence 7, Appli
71	148.5	29.1	809	4	US-08-991-761A-9	Sequence 9, Appli
72	148	29.0	790	4	US-08-991-761A-13	Sequence 13, Appl
73	147.5	28.9	943	2	US-08-469-537A-107	Sequence 107, App
74	146	28.6	800	2	US-08-469-537A-72	Sequence 72, Appl
75	146	28.6	800	2	US-08-469-537A-78	Sequence 78, Appl

#### ALIGNMENTS

RESULT 1  
US-08-797-689-12  
; Sequence 12, Application US/08797689  
; Patent No. 5876969

GENERAL INFORMATION:  
; APPLICANT: Fleer, Reinhard  
; APPLICANT: Fournier, Alain  
; APPLICANT: Guitton, Jean-Dominique  
; APPLICANT: Jung, Gerard

TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
; CONTAINING SAID POLYPEPTIDES  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
; STREET: 500 Arcola Road, 3C43  
; CITY: Collegeville  
; STATE: PA  
; COUNTRY: USA

Search completed: May 25, 2004, 15:03:40  
Job time : 17.0421 secs

```
Sequence 10, Application US/10424999
Publication No. US20040052810A1
NEURAL INFORMATION:
APPLICANT: Nesbit, Mark
APPLICANT: Cameron, Beatrice
APPLICANT: Blanche, Francis
TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
TITLE OF INVENTION: Using Them to Inhibit Angiogenesis
FILE REFERENCE: ST01027-A
CURRENT APPLICATION NUMBER: US/10/424,999
PRIOR FILING DATE: 2003-04-29
PRIOR APPLICATION NUMBER: 10/233,675
PRIOR FILING DATE: 2002-09-04
NUMBER OF SEQ ID NOS: 70
SOFTWARE: PatentIn version 3.2
Q ID NO 10
LENGTH: 87
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Human abrogen as secreted from pBA140 (abrogen N43)
U: 0-424-999-10

Very Match 96.9%; Score 494; DB 12; Length 87;
1st Local Similarity 98.8%; Pred. No. 3.2e-50;
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Q: 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKKNYCRNPDN 60
D: |||||
2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKKNYCRNPDN 61
Q: 61 RRRPWCYVQVGLKPLVQECWVHDCAD 86
D: |||||
62 RRRPWCYVQVGLKPLVQECWVHDCAD 87

RI LT 50
U: 0-425-000-30
Sequence 30, Application US/10425000
Publication No. US20040052777A1
NEURAL INFORMATION:
APPLICANT: Nesbit, Mark
APPLICANT: Cameron, Beatrice
APPLICANT: Blanche, Francis
TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit
TITLE OF INVENTION: Angiogenesis
FILE REFERENCE: ST01027-B
CURRENT APPLICATION NUMBER: US/10/425,000
PRIOR FILING DATE: 2003-04-29
PRIOR APPLICATION NUMBER: 10/233,675
PRIOR FILING DATE: 2002-09-04
NUMBER OF SEQ ID NOS: 105
SOFTWARE: PatentIn version 3.2
Q ID NO 30
LENGTH: 87
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Human derived fusion protein (N43)
U: 0-425-000-30

Very Match 96.9%; Score 494; DB 12; Length 87;
1st Local Similarity 98.8%; Pred. No. 3.2e-50;
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Q: 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKKNYCRNPDN 60
D: |||||
2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKKNYCRNPDN 61
Q: 61 RRRPWCYVQVGLKPLVQECWVHDCAD 86
D: |||||
62 RRRPWCYVQVGLKPLVQECWVHDCAD 87
```

```
/ APPLICANT: Cameron, Beatrice
/ APPLICANT: Blanche, Francis
/ TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
/ TITLE OF INVENTION: Using Them to Inhibit Angiogenesis
/ FILE REFERENCE: ST01027-A
/ CURRENT APPLICATION NUMBER: US/10/424,999
/ PRIOR FILING DATE: 2003-04-29
/ PRIOR APPLICATION NUMBER: 10/233,675
/ PRIOR FILING DATE: 2002-09-04
/ NUMBER OF SEQ ID NOS: 70
/ SOFTWARE: Patent in version 3.2
/ SEQ ID NO 36
/ LENGTH: 221
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Trxa-Abrogen D43 fusion protein
US-10-424-999-36

Query Match          97.8%; Score 499; DB 12; Length 221;
Best Local Similarity 100.0%; Pred. No. 2.3e-50;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

C 1 KTCYEGNGHFGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDN 60
D 136 KTCYEGNGHFGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDN 195
Q 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
D 196 RRRPWCYVQVGLKPLVQECMVHDCAD 221

R JLT 46
U 10-360-101-266
/ Sequence 266, Application US/10360101
/ Publication No. US20040009550A1
/ GENERAL INFORMATION:
/ APPLICANT: Moll, Gert N.
/ APPLICANT: Leenhouts, Cornelis J.
/ TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way
/ FILE REFERENCE: 2183-5673
/ CURRENT APPLICATION NUMBER: US/10/360,101
/ PRIOR FILING DATE: 2003-02-07
/ PRIOR APPLICATION NUMBER: EP 02077060.8
/ PRIOR FILING DATE: 2002-05-24
/ NUMBER OF SEQ ID NOS: 309
/ SOFTWARE: Patent in version 3.1
/ SEQ ID NO 266
/ LENGTH: 445
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: sequence of urokinase
U 10-360-101-266

Query Match          97.8%; Score 499; DB 15; Length 445;
Best Local Similarity 97.8%; Pred. No. 4.9e-50;
Matches 88; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

C 1 KTCYEGNGHFGKASTDTMGRPCLPWNSATVLOQTY--HAHRSALQGLGKHNYCRNP 58
D 70 KTCYEGNGHFGKASTDTMGRPCLPWNSATVLOQTYFTHAHRSDALQGLGKHNYCRNP 129
Q 59 DNRPPWCYVQVGLKPLVQECMVHDCADGK 88
D 130 DNRPPWCYVQVGLKPLVQECMVHDCADGK 159

F JLT 47
U 10-424-999-1
/ Sequence 1, Application US/10424999
/ Publication No. US20040052810A1
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Nesbit, Mark
/ APPLICANT: Cameron, Beatrice
/ APPLICANT: Blanche, Francis
/ TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
/ TITLE OF INVENTION: Using Them to Inhibit Angiogenesis
/ FILE REFERENCE: ST01027-A
/ CURRENT APPLICATION NUMBER: US/10/424,999
/ PRIOR FILING DATE: 2003-04-29
/ PRIOR APPLICATION NUMBER: 10/233,675
/ PRIOR FILING DATE: 2002-09-04
/ NUMBER OF SEQ ID NOS: 70
/ SOFTWARE: Patent in version 3.2
/ SEQ ID NO 1
/ LENGTH: 86
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Human abrogen N43
US-10-424-999-1

Query Match          96.9%; Score 494; DB 12; Length 86;
Best Local Similarity 98.8%; Pred. No. 3.2e-50;
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDN 60
DB 1 KTCYEGNGHFGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDN 60
QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
DB 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86

RESULT 48
US-10-233-675A-1
/ Sequence 1, Application US/10233675A
/ Publication No. US20030228298A1
/ GENERAL INFORMATION:
/ APPLICANT: Nesbit, Mark
/ APPLICANT: Fong, Timothy
/ APPLICANT: Brockstedt, Dirk
/ TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
/ TITLE OF INVENTION: Them To Inhibit Angiogenesis
/ FILE REFERENCE: ST01027
/ CURRENT APPLICATION NUMBER: US/10/233,675A
/ PRIOR FILING DATE: 2002-09-04
/ PRIOR APPLICATION NUMBER: 60/316,300
/ PRIOR FILING DATE: 2001-09-04
/ NUMBER OF SEQ ID NOS: 27
/ SOFTWARE: Patent in version 3.1
/ SEQ ID NO 1
/ LENGTH: 86
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: human derived abrogen
US-10-233-675A-1

Query Match          96.9%; Score 494; DB 15; Length 86;
Best Local Similarity 98.8%; Pred. No. 3.2e-50;
Matches 85; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDN 60
DB 1 KTCYEGNGHFGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDN 60
QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
DB 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86

RESULT 49
US-10-424-999-10
```

```

;
;   URRENT FILING DATE: 2003-04-29
;   RIOR APPLICATION NUMBER: 10/233,675
;   RIOR FILING DATE: 2002-09-04
;   UMBER OF SEQ ID NOS: 70
;   'FTWARE: PatentIn version 3.2
;   Q ID NO 9
;   LENGTH: 87
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Human abrogen as secreted from pmb063 (abrogen D43)
;   0-424-999-9
;
;   ery Match          97.8%; Score 499; DB 12; Length 87;
;   st Local Similarity 100.0%; Pred. No. 8.4e-51;
;   tches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
;   Q: 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
;   D: 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61
;
;   Q: 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
;   D: 62 RRRPWCYVQVGLKPLVQECMVHDCAD 87
;
;   LT 42
;   U: 0-425-000-29
;   ;   quence 29; Application US/10425000
;   ;   blication No. US20040052777A1
;   ;   NERAL INFORMATION:
;   ;   PPLICANT: Nesbit, Mark
;   ;   PPLICANT: Cameron, Beatrice
;   ;   PPLICANT: Blanche, Francis
;   ;   TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit
;   ;   TLE REFERENCE: ST01027-B
;   ;   URRENT APPLICATION NUMBER: US/10/425,000
;   ;   URRENT FILING DATE: 2003-04-29
;   ;   RIOR APPLICATION NUMBER: 10/233,675
;   ;   RIOR FILING DATE: 2002-09-04
;   ;   UMBER OF SEQ ID NOS: 105
;   ;   'FTWARE: PatentIn version 3.2
;   ;   Q ID NO 29
;   ;   LENGTH: 87
;   ;   TYPE: PRT
;   ;   ORGANISM: Artificial Sequence
;   ;   FEATURE:
;   ;   OTHER INFORMATION: Human abrogen (D43)
;   ;   0-425-000-29
;
;   ery Match          97.8%; Score 499; DB 12; Length 87;
;   st Local Similarity 100.0%; Pred. No. 8.4e-51;
;   tches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
;   Q: 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
;   D: 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61
;
;   Q: 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
;   D: 62 RRRPWCYVQVGLKPLVQECMVHDCAD 87
;
;   LT 43
;   U: 0-233-675A-9
;   ;   quence 9; Application US/10233675A
;   ;   blication No. US2003028298A1
;   ;   NERAL INFORMATION:
;   ;   PPLICANT: Nesbit, Mark
;   ;   PPLICANT: Fong, Timothy
;   ;   PPLICANT: Brockstedt, Dirk
;   ;   TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for

```

```

;
;   TITLE OF INVENTION: Them To Inhibit Angiogenesis
;   FILE REFERENCE: ST01027
;   CURRENT APPLICATION NUMBER: US/10/233.675A
;   CURRENT FILING DATE: 2002-09-04
;   PRIOR APPLICATION NUMBER: 60/316,300
;   PRIOR FILING DATE: 2001-09-04
;   NUMBER OF SEQ ID NOS: 27
;   SOFTWARE: PatentIn version 3.1
;   SEQ ID NO 9
;   LENGTH: 87
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: human derived fusion protein
;   US-10-233-675A-9
;
;   Query Match          97.8%; Score 499; DB 15; Length 87;
;   Best Local Similarity 100.0%; Pred. No. 8.4e-51;
;   Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
;   QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
;   Db 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61
;
;   QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
;   Db 62 RRRPWCYVQVGLKPLVQECMVHDCAD 87
;
;   RESULT 44
;   US-10-424-999-37
;   ;   Sequence 37; Application US/10424999
;   ;   Publication No. US20040052810A1
;   ;   GENERAL INFORMATION:
;   ;   APPLICANT: Nesbit, Mark
;   ;   APPLICANT: Cameron, Beatrice
;   ;   APPLICANT: Blanche, Francis
;   ;   TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods fo
;   ;   FILE REFERENCE: ST01027-A
;   ;   CURRENT APPLICATION NUMBER: US/10/424,999
;   ;   CURRENT FILING DATE: 2003-04-29
;   ;   PRIOR APPLICATION NUMBER: 10/233,675
;   ;   PRIOR FILING DATE: 2002-09-04
;   ;   NUMBER OF SEQ ID NOS: 70
;   ;   SOFTWARE: PatentIn version 3.2
;   ;   SEQ ID NO 37
;   ;   LENGTH: 91
;   ;   TYPE: PRT
;   ;   ORGANISM: Artificial Sequence
;   ;   FEATURE:
;   ;   OTHER INFORMATION: Abrogen D43
;   ;   US-10-424-999-37
;
;   Query Match          97.8%; Score 499; DB 12; Length 91;
;   Best Local Similarity 100.0%; Pred. No. 8.8e-51;
;   Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
;   QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
;   Db 6 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 65
;
;   QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
;   Db 66 RRRPWCYVQVGLKPLVQECMVHDCAD 91
;
;   RESULT 45
;   US-10-424-999-36
;   ;   Sequence 36; Application US/10424999
;   ;   Publication No. US20040052810A1
;   ;   GENERAL INFORMATION:
;   ;   APPLICANT: Nesbit, Mark

```



```

; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human kringle domain ATF-Kringle (Abrogen)
U 10-424-999-62

Query Match          97.8%; Score 499; DB 12; Length 86;
Best Local Similarity 100.0%; Pred. No. 8.3e-51;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
D 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Q 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
D 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86

R JLT 38
U 10-425-000-97
; Sequence 97, Application US/10425000
; Publication No. US20040052777A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 97
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human kringle domain ATF-Kringle (Abrogen)
U 10-425-000-97

Query Match          97.8%; Score 499; DB 12; Length 86;
Best Local Similarity 100.0%; Pred. No. 8.3e-51;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
D 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Q 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
D 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86

R JLT 39
U 10-233-675A-5
; Sequence 5, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; PRIOR FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
```

```

; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human derived abrogen
U 10-233-675A-5

Query Match          97.8%; Score 499; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 8.3e-51;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
DB 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86

RESULT 40
U 10-233-675A-22
; Sequence 22, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods fo
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; PRIOR FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fragment of human urokinase plasminogen activator
U 10-233-675A-22

Query Match          97.8%; Score 499; DB 15; Length 86;
Best Local Similarity 100.0%; Pred. No. 8.3e-51;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
QY 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
DB 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86

RESULT 41
U 10-424-999-9
; Sequence 9, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods fo
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
```

```

; Query Match 98.0%; Score 500; DB 12; Length 687;
; 1st Local Similarity 98.9%; Pred. No. 6e-50;
; Matches 86; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Q 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNP 60
D 2 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNP 61

Q 61 RRRPWCYVQVGLKPLVQECMVHDCADG 87
D 62 RRRPWCYVQVGLKPLVQECMVHDCADG 88

; ILT 34
; US-0-425-000-37
; Sequence 37, Application US/10425000
; Publication No. US20040052777A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Kingle Polypeptides and Methods for Using Them to Inhibit
; TITLE OF INVENTION: Angiogenesis
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 37
; LENGTH: 687
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human derived fusion protein
; US-0-425-000-37

; Query Match 98.0%; Score 500; DB 12; Length 687;
; 1st Local Similarity 98.9%; Pred. No. 6e-50;
; Matches 86; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Q 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNP 60
D 2 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNP 61

Q 61 RRRPWCYVQVGLKPLVQECMVHDCADG 87
D 62 RRRPWCYVQVGLKPLVQECMVHDCADG 88

; ILT 35
; US-0-233-675A-17
; Sequence 17, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Them To Inhibit Angiogenesis
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; PRIOR FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 17
; LENGTH: 687
; TYPE: PRT
; ORGANISM: Artificial Sequence
```

```

; FEATURE:
; OTHER INFORMATION: fusion protein human abrogen
; US-10-233-675A-17

; Query Match 98.0%; Score 500; DB 15; Length 687;
; 1st Local Similarity 98.9%; Pred. No. 6e-50;
; Matches 86; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Q 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNP 60
D 2 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNP 61

Q 61 RRRPWCYVQVGLKPLVQECMVHDCADG 87
D 62 RRRPWCYVQVGLKPLVQECMVHDCADG 88

; RESULT 36
; US-10-424-999-5
; Sequence 5, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Using Them to Inhibit Angiogenesis
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 5
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human abrogen
; US-10-424-999-5

; Query Match 97.8%; Score 499; DB 12; Length 86;
; 1st Local Similarity 100.0%; Pred. No. 8.3e-51;
; Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNP 60
D 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNP 60

Q 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86
D 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86

; RESULT 37
; US-10-424-999-62
; Sequence 62, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Using Them to Inhibit Angiogenesis
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 62
```

```

; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 60/323,784
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; Q ID NO 149
; Q ID NO 149
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc.feature
; FEATURE:
; OTHER INFORMATION: Incyte ID No. US20030194721a1 1453334CD1
; US-10-247-671-149

; Query Match
; Best Local Similarity 99.6%; Score 508; DB 14; Length 431;
; Matches 87; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Q 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 60
D 68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 127
Q 61 RRRPWCYVQVGLKPLVQSCMVHDCADGK 88
D 128 RRRPWCYVQVGLKPLVQSCMVHDCADGK 155

; JLT 31
; Sequence 562, Application US/10282174
; Publication No. US20030224380A1
; GENERAL INFORMATION:
; APPLICANT: Becker, Kenneth David
; APPLICANT: Velicelebi, Gonul
; APPLICANT: Elliot, Kathryn J.
; APPLICANT: Wang, Xin
; APPLICANT: Tanzi, Rudolph E.
; APPLICANT: Bertram, Lars
; APPLICANT: Saunders, Aleister J.
; APPLICANT: Mullin, Kristina W.
; APPLICANT: Sampson, Andrew Johnson
; APPLICANT: Blacker, Deborah Lynne
; TITLE OF INVENTION: GENES AND POLYMORPHISMS ON CHROMOSOME 10
; TITLE OF INVENTION: ASSOCIATED WITH ALZHEIMER'S DISEASE AND OTHER
; TITLE OF INVENTION: NEURODEGENERATIVE DISEASES
; FILE REFERENCE: 37481-3308
; CURRENT APPLICATION NUMBER: US/10/282,174
; PRIOR FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: US 60/339,525
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 60/338,010
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/336,929
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/338,363
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 60/337,052
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: US 60/368,919
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 564
; SOFTWARE: FastSeq for Windows Version 4.0
; Q ID NO 562
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: VARIANT
; LOCATION: 15, 58, 141, 214, 231, 274, 366
; OTHER INFORMATION: Xaa = Any Amino Acid
; US-10-282-174-562

; Query Match
; Best Local Similarity 98.2%; Score 501; DB 12; Length 431;

```

```

; Best Local Similarity 98.9%; Pred. No. 2.8e-50;
; Matches 87; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Q 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 60
D 68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 127
Q 61 RRRPWCYVQVGLKPLVQSCMVHDCADGK 88
D 128 RRRPWCYVQVGLKPLVQSCMVHDCADGK 155

; RESULT 32
; US-10-407-821-2
; Sequence 2, Application US/10407821
; Publication No. US20030219386A1
; GENERAL INFORMATION:
; APPLICANT: IDELL, STEVEN
; TITLE OF INVENTION: INTRAPLEURAL SINGLE-CHAIN UROKINASE ALONE OR COMPLEXED
; TITLE OF INVENTION: TO ITS SOLUBLE RECEPTOR PROTECTS AGAINST PLEURAL
; TITLE OF INVENTION: ADHESIONS IN TETRACYCLINE-INDUCED PLEURITIS IN RABBITS
; FILE REFERENCE: UTSN:022US
; CURRENT APPLICATION NUMBER: US/10/407,821
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 60/414,202
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 60/370,466
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-407-821-2

; Query Match
; Best Local Similarity 98.0%; Score 500; DB 15; Length 411;
; Matches 87; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Q 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 60
D 48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 107
Q 61 RRRPWCYVQVGLKPLVQSCMVHDCADGK 88
D 108 RRRPWCYVQVGLKPLVQSCMVHDCADGK 135

; RESULT 33
; US-10-424-999-17
; Sequence 17, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanchet, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods fo
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17
; LENGTH: 687
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein human abrogen
; US-10-424-999-17

```

```

; APPLICANT: Zopf, David
; APPLICANT: Bayer, Robert
; APPLICANT: Hakes, David
; APPLICANT: Chen, Xi
; APPLICANT: Bove, Caryn
; TITLE OF INVENTION: GRANULOCYTE COLONY STIMULATING FACTOR: REMODELING AND
; TITLE OF INVENTION: GLYCOCONJUGATION OF G-CSF
; FILE REFERENCE: 040853-01-5054
; URGENT APPLICATION NUMBER: US/10/410,962
; URGENT FILING DATE: 2003-04-09
; RIOR APPLICATION NUMBER: US 60/328,523
; RIOR FILING DATE: 2001-10-10
; RIOR APPLICATION NUMBER: US 60/344,692
; RIOR FILING DATE: 2001-10-19
; RIOR APPLICATION NUMBER: US 60/387,292
; RIOR FILING DATE: 2002-06-07
; RIOR APPLICATION NUMBER: US 60/391,777
; RIOR FILING DATE: 2002-06-25
; RIOR APPLICATION NUMBER: US 60/396,594
; RIOR FILING DATE: 2002-07-17
; RIOR APPLICATION NUMBER: US 60/404,249
; RIOR FILING DATE: 2002-08-16
; RIOR APPLICATION NUMBER: US 60/407,527
; RIOR FILING DATE: 2002-08-28
; UMBER OF SEQ ID NOS: 75
; SOFTWARE: Patent in version 3.2
; Q ID NO 34
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; 0-410-962-34

ery Match 100.0%; Score 510; DB 16; Length 431;
st Local Similarity 100.0%; Pred. No. 2.4e-51;
tches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q: 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNPDN 60
DI 68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNPDN 127
Q: 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
DI 128 RRRPWCYVQVGLKPLVQECMVHDCADGK 155

RI LT 28
UR 0-411-049-34
; quence 34, Application US/10411049
; blication No. US20040082026A1
; NERAL INFORMATION:
; PPLICANT: Neose Technologies, Inc.
; PPLICANT: Defrees, Shawn
; PPLICANT: Zopf, David
; PPLICANT: Bayer, Robert
; PPLICANT: Hakes, David
; PPLICANT: Chen, Xi
; PPLICANT: Bove, Caryn
; TITLE OF INVENTION: INTERFERON ALPHA: REMODELING AND GLYCOCONJUGATION OF INTERFERON
; TITLE OF INVENTION: ALPHA
; URGENT APPLICATION NUMBER: US/10/411,049
; URGENT FILING DATE: 2003-04-09
; RIOR APPLICATION NUMBER: US 60/328,523
; RIOR FILING DATE: 2001-10-10
; RIOR APPLICATION NUMBER: US 60/344,692
; RIOR FILING DATE: 2001-10-19
; RIOR APPLICATION NUMBER: US 60/387,292
; RIOR FILING DATE: 2002-06-07
; RIOR APPLICATION NUMBER: US 60/391,777
; RIOR FILING DATE: 2002-06-25
; RIOR APPLICATION NUMBER: US 60/396,594
; RIOR FILING DATE: 2002-07-17
; RIOR APPLICATION NUMBER: US 60/404,249

; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: US 60/407,527
; PRIOR FILING DATE: 2002-08-28
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 34
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-411-049-34

Query Match 100.0%; Score 510; DB 16; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.4e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNPDN 60
DB 68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNPDN 127
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 128 RRRPWCYVQVGLKPLVQECMVHDCADGK 155

RESULT 29
US-10-087-192-594
; Sequence 594, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 594
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-087-192-594

Query Match 100.0%; Score 510; DB 12; Length 437;
Best Local Similarity 100.0%; Pred. No. 2.5e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNPDN 60
DB 74 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNPDN 133
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 134 RRRPWCYVQVGLKPLVQECMVHDCADGK 161

RESULT 30
US-10-247-671-149
; Sequence 149, Application US/10247671
; Publication No. US20030194721A1
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247,671
```

US 10-131-985-21  
Query Match 100.0%; Score 510; DB 14; Length 431;  
Best Local Similarity 100.0%; Pred. No. 2.4e-51;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 60  
D 68 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 127

Q 61 RRRPWCYVQVGLKPLVQECMWHDCADGK 88  
D 128 RRRPWCYVQVGLKPLVQECMWHDCADGK 155

R JLT 25  
US 10-295-027-414  
Query Match 100.0%; Score 510; DB 15; Length 431;  
Best Local Similarity 100.0%; Pred. No. 2.4e-51;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 60  
D 68 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 127

Q 61 RRRPWCYVQVGLKPLVQECMWHDCADGK 88  
D 128 RRRPWCYVQVGLKPLVQECMWHDCADGK 155

R JLT 25  
US 10-295-027-414  
Query Match 100.0%; Score 510; DB 15; Length 431;  
Best Local Similarity 100.0%; Pred. No. 2.4e-51;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 60  
D 68 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 127

Q 61 RRRPWCYVQVGLKPLVQECMWHDCADGK 88  
D 128 RRRPWCYVQVGLKPLVQECMWHDCADGK 155

US 10-295-027-1275  
Query Match 100.0%; Score 510; DB 15; Length 431;  
Best Local Similarity 100.0%; Pred. No. 2.4e-51;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 60  
D 68 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 127

Q 61 RRRPWCYVQVGLKPLVQECMWHDCADGK 88  
D 128 RRRPWCYVQVGLKPLVQECMWHDCADGK 155

R JLT 25  
US 10-295-027-1275  
Query Match 100.0%; Score 510; DB 15; Length 431;  
Best Local Similarity 100.0%; Pred. No. 2.4e-51;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 60  
D 68 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 127

Q 61 RRRPWCYVQVGLKPLVQECMWHDCADGK 88  
D 128 RRRPWCYVQVGLKPLVQECMWHDCADGK 155

R JLT 25  
US 10-295-027-1275  
Query Match 100.0%; Score 510; DB 15; Length 431;  
Best Local Similarity 100.0%; Pred. No. 2.4e-51;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 60  
D 68 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 127

Q 61 RRRPWCYVQVGLKPLVQECMWHDCADGK 88  
D 128 RRRPWCYVQVGLKPLVQECMWHDCADGK 155

R JLT 25  
US 10-295-027-1275  
Query Match 100.0%; Score 510; DB 15; Length 431;  
Best Local Similarity 100.0%; Pred. No. 2.4e-51;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 60  
D 68 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 127

Q 61 RRRPWCYVQVGLKPLVQECMWHDCADGK 88  
D 128 RRRPWCYVQVGLKPLVQECMWHDCADGK 155

```

; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; PRIOR FILING DATE: 2002-06-12
; RIOR FILING DATE: 2001-06-13
; RIOR FILING DATE: 2001-06-13
; RIOR FILING DATE: 2001-06-13
; RIOR FILING DATE: 2001-06-13
; RIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; Q ID NO 184
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; 0-171-311-184

ery Match 100.0%; Score 510; DB 14; Length 431;
st Local Similarity 100.0%; Pred. No. 2.4e-51;
tches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q: 1 KTCYEGNGHFYRGKASTDTMGRCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 60
D: 68 KTCYEGNGHFYRGKASTDTMGRCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 127
Q: 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
D: 128 RRRPWCYVQVGLKPLVQECMVHDCADGK 155

LT 22
U: 0-193-656-4
; quence 4, Application US/10/193656
; blication No. US20030096733A1
; NERAL INFORMATION:
; PPLICANT: NY, Tor
; PPLICANT: HOLMDAHL, Rikard
; PPLICANT: LI, Jinan
; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
; FILE REFERENCE: 3810/1J577-US3
; CURRENT APPLICATION NUMBER: US/10/193,656
; PRIOR FILING DATE: 2002-07-10
; RIOR APPLICATION NUMBER: US 60/304,461
; RIOR FILING DATE: 2001-07-10
; RIOR APPLICATION NUMBER: US 60/304,490
; RIOR FILING DATE: 2001-07-10
; RIOR APPLICATION NUMBER: US 60/305,182
; RIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; Q ID NO 4
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; UBICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank / P00749
; DATABASE ENTRY DATE: 1986-07-21
; RELEVANT RESIDUES: (1)...(431)
; 0-193-656-4

ery Match 100.0%; Score 510; DB 14; Length 431;
st Local Similarity 100.0%; Pred. No. 2.4e-51;
tches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q: 1 KTCYEGNGHFYRGKASTDTMGRCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 60
D: 68 KTCYEGNGHFYRGKASTDTMGRCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 127
Q: 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
D: 128 RRRPWCYVQVGLKPLVQECMVHDCADGK 155

US-10-301-822-161
; Sequence 161, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Bursart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MPM01-029P2RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 161
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-10-301-822-161

Query Match 100.0%; Score 510; DB 14; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.4e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 60
Db 68 KTCYEGNGHFYRGKASTDTMGRCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 127
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADGK 155

RESULT 24
US-10-131-985-21
; Sequence 21, Application US/10131985
; Publication No. US20030199440A1
; GENERAL INFORMATION:
; APPLICANT: Dack, Kevin N
; APPLICANT: Davies, Michael J
; APPLICANT: Fish, Paul V
; APPLICANT: Huggins, Jonathan P
; APPLICANT: McIntosh, Fraser S
; APPLICANT: Occleston, Nicholas L
; TITLE OF INVENTION: Composition
; FILE REFERENCE: PCS 10391A
; CURRENT APPLICATION NUMBER: US/10/131,985
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: US/09/726,295
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: GB 9930768.8
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens

```

APPLICANT: Zopf, David  
APPLICANT: Bayer, Robert  
APPLICANT: Hakes, David  
APPLICANT: Chen, Xi  
APPLICANT: Bove, Caryn  
TITLE OF INVENTION: ALPHA GALACTOSIDASE A: REMODELING AND GLYCOCONJUGATION OF ALPHA  
TITLE OF INVENTION: GALACTOSIDASE A  
CURRENT APPLICATION NUMBER: US/10/411,037  
CURRENT FILING DATE: 2003-04-09  
PRIOR APPLICATION NUMBER: US 60/328,523  
PRIOR FILING DATE: 2001-10-10  
PRIOR APPLICATION NUMBER: US 60/344,692  
PRIOR FILING DATE: 2001-10-19  
PRIOR APPLICATION NUMBER: US 60/387,292  
PRIOR FILING DATE: 2002-06-07  
PRIOR APPLICATION NUMBER: US 60/391,777  
PRIOR FILING DATE: 2002-06-25  
PRIOR APPLICATION NUMBER: US 60/396,594  
PRIOR FILING DATE: 2002-07-17  
PRIOR APPLICATION NUMBER: US 60/404,249  
PRIOR FILING DATE: 2002-08-16  
PRIOR APPLICATION NUMBER: US 60/407,527  
PRIOR FILING DATE: 2002-08-28  
NUMBER OF SEQ ID NOS: 75  
SOFTWARE: Patent in version 3.2  
SEQ ID NO 34  
LENGTH: 431  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-037-34

Query Match 100.0%; Score 510; DB 12; Length 431;  
Best Local Similarity 100.0%; Pred. No. 2.4e-51;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q: 1 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 60  
D: 68 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 127

Q: 61 RRRPWCYVQVGLKPLVQECMWHDCADGK 88  
D: 128 RRRPWCYVQVGLKPLVQECMWHDCADGK 155

LT 19

US-10-026-34  
Sequence 34, Application US/10411026  
Publication No. US20040063911A1  
GENERAL INFORMATION:  
APPLICANT: Neose Technologies, Inc.  
APPLICANT: Zopf, David  
APPLICANT: Bayer, Robert  
APPLICANT: Hakes, David  
APPLICANT: Chen, Xi  
TITLE OF INVENTION: PROTEIN REMODELING METHODS AND PEPTIDES PRODUCED BY THE  
TITLE OF INVENTION: METHODS  
CURRENT APPLICATION NUMBER: US/10/411,026  
CURRENT FILING DATE: 2003-04-09  
PRIOR APPLICATION NUMBER: US 60/328,523  
PRIOR FILING DATE: 2001-10-10  
PRIOR APPLICATION NUMBER: US 60/344,692  
PRIOR FILING DATE: 2001-10-19  
PRIOR APPLICATION NUMBER: US 60/387,292  
PRIOR FILING DATE: 2002-06-07  
PRIOR APPLICATION NUMBER: US 60/391,777  
PRIOR FILING DATE: 2002-06-25  
PRIOR APPLICATION NUMBER: US 60/396,594  
PRIOR FILING DATE: 2002-07-17  
PRIOR APPLICATION NUMBER: US 60/404,249  
PRIOR FILING DATE: 2002-08-16

PRIOR APPLICATION NUMBER: US 60/407,527  
PRIOR FILING DATE: 2002-08-28  
NUMBER OF SEQ ID NOS: 75  
SOFTWARE: Patent in version 3.2  
SEQ ID NO 34  
LENGTH: 431  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-411-026-34

Query Match 100.0%; Score 510; DB 12; Length 431;  
Best Local Similarity 100.0%; Pred. No. 2.4e-51;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 60  
DB 68 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 127

QY 61 RRRPWCYVQVGLKPLVQECMWHDCADGK 88  
DB 128 RRRPWCYVQVGLKPLVQECMWHDCADGK 155

RESULT 20  
US-10-076-421-2  
Sequence 2, Application US/10076421  
Publication No. US20020193304A1  
GENERAL INFORMATION:  
APPLICANT: WADA, MANABU  
APPLICANT: WADA, NAOKO  
TITLE OF INVENTION: ANTI-HIV AGENTS  
FILE REFERENCE: HAYAK-9  
CURRENT APPLICATION NUMBER: US/10/076,421  
CURRENT FILING DATE: 2002-05-17  
PRIOR APPLICATION NUMBER: JP 2001-42655  
PRIOR FILING DATE: 2001-02-20  
PRIOR APPLICATION NUMBER: JP 2001-184284  
PRIOR FILING DATE: 2001-06-19  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: Patent in Ver. 2.1  
SEQ ID NO 2  
LENGTH: 431  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-076-421-2

Query Match 100.0%; Score 510; DB 13; Length 431;  
Best Local Similarity 100.0%; Pred. No. 2.4e-51;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 60  
DB 68 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 127

QY 61 RRRPWCYVQVGLKPLVQECMWHDCADGK 88  
DB 128 RRRPWCYVQVGLKPLVQECMWHDCADGK 155

RESULT 21  
US-10-171-311-184  
Sequence 184, Application US/10171311  
Publication No. US20030087270A1  
GENERAL INFORMATION:  
APPLICANT: Schlegel, Robert  
APPLICANT: Chen, Yan  
APPLICANT: Zhao, Xumei  
APPLICANT: Monahan, John  
APPLICANT: Kamatkar, Shubhangi  
APPLICANT: Glatt, Karen  
APPLICANT: Gannavarapu, Manjula  
APPLICANT: Hoeresh, Sebastian  
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR

```

; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
; 0-264-049-2927

Query Match
Best Local Similarity 100.0%; Score 510; DB 15; Length 337;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLTQTYHAHRSDALQGLGKHNCRNPDN 60
DB 74 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLTQTYHAHRSDALQGLGKHNCRNPDN 133
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 134 RRRPWCYVQVGLKPLVQECMVHDCADGK 161

RESULT 15
US-09-880-503-6
; Sequence 6, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; URGENT APPLICATION NUMBER: US/09/880,503
; URGENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; Q ID NO 6
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
; 9-880-503-6

Query Match
Best Local Similarity 100.0%; Score 510; DB 9; Length 403;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLTQTYHAHRSDALQGLGKHNCRNPDN 60
DB 48 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLTQTYHAHRSDALQGLGKHNCRNPDN 107
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

RESULT 16
US-09-880-503-3
; Sequence 3, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; URGENT APPLICATION NUMBER: US/09/880,503
; URGENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; Q ID NO 3
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
; 9-880-503-3
```

```

US-09-880-503-3
Query Match
Best Local Similarity 100.0%; Score 510; DB 9; Length 411;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLTQTYHAHRSDALQGLGKHNCRNPDN 60
DB 48 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLTQTYHAHRSDALQGLGKHNCRNPDN 107
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

RESULT 17
US-09-264-468B-1
; Sequence 1, Application US/09264468B
; Patent No. US20020106775A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Jieyi
; APPLICANT: Nienaber, Vicki L.
; APPLICANT: Henkin, Jack
; APPLICANT: Smith, Richard A.
; APPLICANT: Walter, Karl A.
; APPLICANT: Severin, Jean M.
; APPLICANT: Edalji, Rohinton
; APPLICANT: Johnson Jr., Robert W.
; APPLICANT: Holzman, Thomas F.
; TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE
; FILE REFERENCE: 6310.US.P1
; CURRENT APPLICATION NUMBER: US/09/264,468B
; URGENT FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: US 09/036,361
; PRIOR FILING DATE: 1998-03-06
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(20)
; OTHER INFORMATION: Leader sequence
; NAME/KEY: VARIANT
; LOCATION: (279)...(279)
; OTHER INFORMATION: Xaa = any amino acid
; NAME/KEY: VARIANT
; LOCATION: (302)...(302)
; OTHER INFORMATION: Xaa = any amino acid
; US-09-264-468B-1

Query Match
Best Local Similarity 100.0%; Score 510; DB 9; Length 431;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLTQTYHAHRSDALQGLGKHNCRNPDN 60
DB 68 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLTQTYHAHRSDALQGLGKHNCRNPDN 127
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 128 RRRPWCYVQVGLKPLVQECMVHDCADGK 155

RESULT 18
US-10-411-037-34
; Sequence 34, Application US/10411037
; Publication No. US2004004346A1
; GENERAL INFORMATION:
; APPLICANT: Neose Technologies, Inc.
; APPLICANT: DePress, Shawn
```



PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, 3c43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: System 7.1  
SOFTWARE: Word 5.1 (PatentIn)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/702,636  
FILING DATE: 06-Nov-2003  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/797,689  
FILING DATE: 31-JAN-1997  
APPLICATION NUMBER: US 08/256,927  
FILING DATE: 28-JUL-1994  
APPLICATION NUMBER: FR 92/01064  
FILING DATE: 31-JAN-1992  
APPLICATION NUMBER: PCT/FR93/00085  
FILING DATE: 28-JAN-1993

ATTORNEY/AGENT INFORMATION:  
NAME: Smith Ph.D., Julie K.  
REGISTRATION NUMBER: P-38,619  
REFERENCE/DOCKET NUMBER: ST92006-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 138 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US 10-702-636-12

Query Match 100.0%; Score 510; DB 16; Length 138;  
Best Local Similarity 100.0%; Pred. No. 7.1e-52;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q 1 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 60  
D 51 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 110

Q 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
D 111 RRRPWCYVQVGLKPLVQECMVHDCADGK 138

RESULT 12  
US-10-880-503-8  
Sequence 8, Application US/09880503  
Publication No. US20020131964A1  
GENERAL INFORMATION:  
APPLICANT: CINES, Douglas B  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND  
FILE REFERENCE: 9596-331  
CURRENT APPLICATION NUMBER: US/09/880,503  
PRIOR FILING DATE: 2001-06-13  
PRIOR APPLICATION NUMBER: US 60/212,847  
PRIOR FILING DATE: 2000-06-20  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 8  
LENGTH: 143  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-880-503-8

Query Match 100.0%; Score 510; DB 9; Length 143;  
Best Local Similarity 100.0%; Pred. No. 7.4e-52;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 60  
DB 48 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 107

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
DB 108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

RESULT 13  
US-10-106-698-6266  
Sequence 6266, Application US/10106698  
Publication No. US20030109690A1  
GENERAL INFORMATION:  
APPLICANT: Ruben et al.  
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptid  
FILE REFERENCE: PA005P1  
CURRENT APPLICATION NUMBER: US/10/106,698  
PRIOR FILING DATE: 2002-03-27  
PRIOR APPLICATION NUMBER: PCT/US00/26524  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US 60/157,137  
PRIOR FILING DATE: 1999-09-29  
PRIOR APPLICATION NUMBER: US 60/163,280  
PRIOR FILING DATE: 1999-11-03  
NUMBER OF SEQ ID NOS: 8564  
SOFTWARE: PatentIn Ver. 3.0  
SEQ ID NO 6266  
LENGTH: 337  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-106-698-6266

Query Match 100.0%; Score 510; DB 14; Length 337;  
Best Local Similarity 100.0%; Pred. No. 1.9e-51;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 60  
DB 74 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 133

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
DB 134 RRRPWCYVQVGLKPLVQECMVHDCADGK 161

RESULT 14  
US-10-264-049-2927  
Sequence 2927, Application US/10264049  
Publication No. US20040005579A1  
GENERAL INFORMATION:  
APPLICANT: Birse et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PA133P1  
CURRENT APPLICATION NUMBER: US/10/264,049  
CURRENT FILING DATE: 2002-10-04  
PRIOR APPLICATION NUMBER: PCT/US01/18569  
PRIOR FILING DATE: 2001-06-07  
PRIOR APPLICATION NUMBER: US 60/209,467  
PRIOR FILING DATE: 2000-06-07  
NUMBER OF SEQ ID NOS: 4360  
SOFTWARE: PatentIn Ver. 3.1  
SEQ ID NO 2927

US 0-237-624-12  
; Sequence 12, Application US/10237624  
; Application No. US20030082747A1  
; GENERAL INFORMATION:  
; APPLICANT: Fleer, Reinhard  
; Fournier, Alain  
; Guillon, Jean-Dominique  
; Jung, Gerard  
; Yeh, Patrice  
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
; CONTAINING SAID POLYPEPTIDES  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rhone-Poulenc Rorer Inc.  
; STREET: 500 Arcola Road, 3C43  
; CITY: Collegeville  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19426  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: Macintosh  
; OPERATING SYSTEM: System 7.1  
; SOFTWARE: Word 5.1 (PatentIn)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/237,624  
; FILING DATE: 10-Sep-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/797,689  
; FILING DATE: 31-JAN-1997  
; APPLICATION NUMBER: US 08/256,927  
; FILING DATE: 28-JUL-1994  
; APPLICATION NUMBER: FR 92/01064  
; FILING DATE: 31-JAN-1992  
; APPLICATION NUMBER: PCT/FR93/00085  
; FILING DATE: 28-JAN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith Ph.D., Julie K.  
; REGISTRATION NUMBER: P-38,619  
; REFERENCE/DOCKET NUMBER: ST92006-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (610) 454-3839  
; TELEFAX: (610) 454-3808  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 138 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US 0-237-624-12  
Query Match 100.0%; Score 510; DB 14; Length 138;  
Best Local Similarity 100.0%; Pred. No. 7.1e-52;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHFGKASDTMGRCPLPWSATVLQOQTYHAHRSALQGLGKHNYCRNPDN 60  
D 51 KTCYEGNGHFGKASDTMGRCPLPWSATVLQOQTYHAHRSALQGLGKHNYCRNPDN 110  
QY 61 RRRPWCYVQGLKPLVQECMVHDCADGK 88  
D 111 RRRPWCYVQGLKPLVQECMVHDCADGK 138  
RESULT 11  
US-10-702-536-12  
; Sequence 12, Application US/10702536  
; Publication No. US20040086976A1  
; GENERAL INFORMATION:  
; APPLICANT: Fleer, Reinhard  
; Fournier, Alain  
; Guillon, Jean-Dominique  
; Jung, Gerard  
; Yeh, Patrice  
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,

Fournier, Alain  
Guillon, Jean-Dominique  
Jung, Gerard  
Yeh, Patrice  
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
CONTAINING SAID POLYPEPTIDES  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: System 7.1  
SOFTWARE: Word 5.1 (PatentIn)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/702,536  
FILING DATE: 07-Nov-2003  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/797,689  
FILING DATE: 31-JAN-1997  
APPLICATION NUMBER: US 08/256,927  
FILING DATE: 28-JUL-1994  
APPLICATION NUMBER: FR 92/01064  
FILING DATE: 31-JAN-1992  
APPLICATION NUMBER: PCT/FR93/00085  
FILING DATE: 28-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith Ph.D., Julie K.  
REGISTRATION NUMBER: P-38,619  
REFERENCE/DOCKET NUMBER: ST92006-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 138 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-10-702-536-12  
Query Match 100.0%; Score 510; DB 16; Length 138;  
Best Local Similarity 100.0%; Pred. No. 7.1e-52;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHFGKASDTMGRCPLPWSATVLQOQTYHAHRSALQGLGKHNYCRNPDN 60  
D 51 KTCYEGNGHFGKASDTMGRCPLPWSATVLQOQTYHAHRSALQGLGKHNYCRNPDN 110  
QY 61 RRRPWCYVQGLKPLVQECMVHDCADGK 88  
D 111 RRRPWCYVQGLKPLVQECMVHDCADGK 138  
RESULT 11  
US-10-702-636-12  
; Sequence 12, Application US/10702636  
; Publication No. US20040086977A1  
; GENERAL INFORMATION:  
; APPLICANT: Fleer, Reinhard  
; Fournier, Alain  
; Guillon, Jean-Dominique  
; Jung, Gerard  
; Yeh, Patrice  
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,

Q 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQGLGKHNYCRNPDN 60  
D 51 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQGLGKHNYCRNPDN 110  
Q 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
D 111 RRRPWCYVQVGLKPLVQECMVHDCADGK 138  
R LT 7  
U 0-237-866-12  
Sequence 12, Application US/10237866  
Publication No. US20030036172A1  
GENERAL INFORMATION:  
APPLICANT: Fleer, Reinhard  
Fournier, Alain  
Guitton, Jean-Dominique  
Jung, Gerard  
Yeh, Patrice  
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
CONTAINING SAID POLYPEPTIDES  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: System 7.1  
SOFTWARE: Word 5.1 (Patentin)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/237,866  
FILING DATE: 10-Sep-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/797,689  
FILING DATE: 31-JAN-1997  
APPLICATION NUMBER: US 08/256,927  
FILING DATE: 28-JUL-1994  
APPLICATION NUMBER: FR 92/01064  
FILING DATE: 31-JAN-1992  
APPLICATION NUMBER: PCT/FR93/00085  
FILING DATE: 28-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith Ph.D., Julie K.  
REGISTRATION NUMBER: P-38,619  
REFERENCE/DOCKET NUMBER: ST92006-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 138 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
U 0-237-866-12  
Query Match 100.0%; Score 510; DB 14; Length 138;  
Best Local Similarity 100.0%; Pred. No. 7.1e-52;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Q 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQGLGKHNYCRNPDN 60  
D 51 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQGLGKHNYCRNPDN 110  
Q 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88

Db 111 RRRPWCYVQVGLKPLVQECMVHDCADGK 138  
RESULT 8  
US-10-237-871-12  
Sequence 12, Application US/10237871  
Publication No. US20030036172A1  
GENERAL INFORMATION:  
APPLICANT: Fleer, Reinhard  
Fournier, Alain  
Guitton, Jean-Dominique  
Jung, Gerard  
Yeh, Patrice  
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
CONTAINING SAID POLYPEPTIDES  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: System 7.1  
SOFTWARE: Word 5.1 (Patentin)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/237,871  
FILING DATE: 10-Sep-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/797,689  
FILING DATE: 31-JAN-1997  
APPLICATION NUMBER: US 08/256,927  
FILING DATE: 28-JUL-1994  
APPLICATION NUMBER: FR 92/01064  
FILING DATE: 31-JAN-1992  
APPLICATION NUMBER: PCT/FR93/00085  
FILING DATE: 28-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith Ph.D., Julie K.  
REGISTRATION NUMBER: P-38,619  
REFERENCE/DOCKET NUMBER: ST92006-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (610) 454-3839  
TELEFAX: (610) 454-3808  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 138 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-10-237-871-12  
Query Match 100.0%; Score 510; DB 14; Length 138;  
Best Local Similarity 100.0%; Pred. No. 7.1e-52;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQGLGKHNYCRNPDN 60  
Db 51 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQGLGKHNYCRNPDN 110  
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
Db 111 RRRPWCYVQVGLKPLVQECMVHDCADGK 138  
RESULT 9

Query Match	100.0%;	Score 510;	DB 14;	Length 138;
Best Local Similarity	100.0%;	Pred. No. 7.1e-52;		
Matches	88;	Conservative 0;	Mismatches 0;	Indels 0;
Gaps	0;			
QY	1	KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHYCRNP DN 60		
DB	51	KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHYCRNP DN 110		
QY	61	RRRPWCYVQVGLKPLVQECMVHDCADGK 88		
DB	111	RRRPWCYVQVGLKPLVQECMVHDCADGK 138		
RESULT 6				
US-10-237-708-12				
Sequence 12, Application US/10237708				
Publication No. US20030036170A1				
GENERAL INFORMATION:				
APPLICANT: Fleer, Reinhard				
Fournier, Alain				
Guitton, Jean-Dominique				
Jung, Gerard				
Yeh, Patrice				
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES, PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION CONTAINING SAID POLYPEPTIDES				
NUMBER OF SEQUENCES: 36				
CORRESPONDENCE ADDRESS:				
ADDRESSEE: Rhone-Poulenc Rorer Inc.				
STREET: 500 Arcola Road, 3C43				
CITY: Collegeville				
STATE: PA				
COUNTRY: USA				
ZIP: 19426				
COMPUTER READABLE FORM:				
MEDIUM TYPE: Floppy disk				
COMPUTER: Macintosh				
OPERATING SYSTEM: System 7.1				
SOFTWARE: Word 5.1 (Patentin)				
CURRENT APPLICATION DATA:				
APPLICATION NUMBER: US/10/237,667				
FILING DATE: 10-Sep-2002				
CLASSIFICATION: <Unknown>				
PRIOR APPLICATION DATA:				
APPLICATION NUMBER: US/08/797,689				
FILING DATE: 31-JAN-1997				
APPLICATION NUMBER: US 08/256,927				
FILING DATE: 28-JUL-1994				
APPLICATION NUMBER: FR 92/01064				
FILING DATE: 31-JAN-1992				
APPLICATION NUMBER: PCT/FR93/00085				
FILING DATE: 28-JAN-1993				
ATTORNEY/AGENT INFORMATION:				
NAME: Smith Ph.D., Julie K.				
REGISTRATION NUMBER: P-38,619				
REFERENCE/DOCKET NUMBER: ST92006-US				
TELECOMMUNICATION INFORMATION:				
TELEPHONE: (610) 454-3839				
TELEFAX: (610) 454-3808				
INFORMATION FOR SEQ ID NO: 12:				
SEQUENCE CHARACTERISTICS:				
LENGTH: 138 amino acids				
TYPE: amino acid				
TOPOLOGY: linear				
MOLECULE TYPE: protein				
SEQUENCE DESCRIPTION: SEQ ID NO: 12:				
US-10-237-708-12				
Query Match				
100.0%; Score 510; DB 14; Length 138;				
Best Local Similarity				
100.0%; Pred. No. 7.1e-52;				
Matches				
88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Query				
1 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHYCRNP DN 60				
Database				
51 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHYCRNP DN 110				
Query				
61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88				
Database				
111 RRRPWCYVQVGLKPLVQECMVHDCADGK 138				
Result 5				
US-10-237-667-12				
Sequence 12, Application US/10237667				
Publication No. US20030022308A1				
GENERAL INFORMATION:				
APPLICANT: Fleer, Reinhard				
Fournier, Alain				
Guitton, Jean-Dominique				
Jung, Gerard				
Yeh, Patrice				
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES, PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION CONTAINING SAID POLYPEPTIDES				
NUMBER OF SEQUENCES: 36				
CORRESPONDENCE ADDRESS:				
ADDRESSEE: Rhone-Poulenc Rorer Inc.				
STREET: 500 Arcola Road, 3C43				
CITY: Collegeville				
STATE: PA				
COUNTRY: USA				
ZIP: 19426				
COMPUTER READABLE FORM:				
MEDIUM TYPE: Floppy disk				
COMPUTER: Macintosh				
OPERATING SYSTEM: System 7.1				
SOFTWARE: Word 5.1 (Patentin)				
CURRENT APPLICATION DATA:				
APPLICATION NUMBER: US/10/237,667				
FILING DATE: 10-Sep-2002				
CLASSIFICATION: <Unknown>				
PRIOR APPLICATION DATA:				
APPLICATION NUMBER: US/08/797,689				
FILING DATE: 31-JAN-1997				
APPLICATION NUMBER: US 08/256,927				
FILING DATE: 28-JUL-1994				
APPLICATION NUMBER: FR 92/01064				
FILING DATE: 31-JAN-1992				
APPLICATION NUMBER: PCT/FR93/00085				
FILING DATE: 28-JAN-1993				
ATTORNEY/AGENT INFORMATION:				
NAME: Smith Ph.D., Julie K.				
REGISTRATION NUMBER: P-38,619				
REFERENCE/DOCKET NUMBER: ST92006-US				
TELECOMMUNICATION INFORMATION:				
TELEPHONE: (610) 454-3839				
TELEFAX: (610) 454-3808				
INFORMATION FOR SEQ ID NO: 12:				
SEQUENCE CHARACTERISTICS:				
LENGTH: 138 amino acids				
TYPE: amino acid				
TOPOLOGY: linear				
MOLECULE TYPE: protein				
SEQUENCE DESCRIPTION: SEQ ID NO: 12:				
US-10-237-667-12				
Query Match				
100.0%; Score 510; DB 14; Length 138;				
Best Local Similarity				
100.0%; Pred. No. 7.1e-52;				
Matches				
88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Query				
1 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHYCRNP DN 60				
Database				
51 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHYCRNP DN 110				
Query				
61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88				
Database				
111 RRRPWCYVQVGLKPLVQECMVHDCADGK 138				
Result 5				
US-10-237-667-12				

```

; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1
; TYPE: PRT
; LENGTH: 88
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
; US-880-503-1
;
; Query Match 100.0%; Score 510; DB 9; Length 88;
; Best Local Similarity 100.0%; Pred. NO. 4.3e-52;
; Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Q1 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDN 60
; D1 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDN 60
; Q1 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
; D1 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
;
; RESULT 4
; US-09-880-503-9
; Sequence 9, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 9
; TYPE: PRT
; LENGTH: 96
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
; US-880-503-9
;
; Query Match 100.0%; Score 510; DB 9; Length 96;
; Best Local Similarity 100.0%; Pred. NO. 4.8e-52;
; Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Q1 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDN 60
; D1 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDN 60
; Q1 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
; D1 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
;
; RESULT 3
; US-880-503-4
; Sequence 4, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
;
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; TYPE: PRT
; LENGTH: 135
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
; US-09-880-503-4
;
; Query Match 100.0%; Score 510; DB 9; Length 135;
; Best Local Similarity 100.0%; Pred. NO. 6.9e-52;
; Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDN 60
; DB 48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDN 107
; QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
; DB 108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135
;
; RESULT 4
; US-09-984-186-12
; Sequence 12, Application US/09984186
; Patent No. US2002015101A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guittou, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (Patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,186
; FILING DATE: 29-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIORITY INFORMATION:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-36,619
; REFERENCE/DOCKET NUMBER: ST92005-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

Object: protein - protein search, using sw model

Run on: May 25, 2004, 14:53:05 ; Search time 16.0421 Seconds  
(without alignments)  
1530.045 Million cell updates/sec

Target: US-09-880-503-1

Percent score: 51.0

Sequence: 1 KTCYEGNGHFYRGKASTDTM.....QVGLKFLVQCMVHDCADGK 88

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Sequences: 1149313 seqs, 278921704 residues

Total number of hits satisfying chosen parameters: 1149313

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database: Published Applications AA:

1:	/cgm2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2:	/cgm2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgm2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4:	/cgm2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5:	/cgm2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6:	/cgm2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgm2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8:	/cgm2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9:	/cgm2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10:	/cgm2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgm2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12:	/cgm2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13:	/cgm2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14:	/cgm2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15:	/cgm2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16:	/cgm2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17:	/cgm2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18:	/cgm2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Score	Query Match	Length	ID	Description
1	51.0	100.0	88	9	US-09-880-503-1
2	51.0	100.0	95	9	US-09-880-503-9
3	51.0	100.0	135	9	US-09-880-503-4
4	51.0	100.0	138	9	US-09-880-503-12
5	51.0	100.0	138	14	US-10-237-667-12
6	51.0	100.0	138	14	US-10-237-708-12
7	51.0	100.0	138	14	US-10-237-866-12
8	51.0	100.0	138	14	US-10-237-871-12
9	51.0	100.0	138	14	US-10-237-624-12
10	51.0	100.0	138	16	US-10-702-536-12
11	51.0	100.0	138	16	US-10-702-636-12
12	51.0	100.0	143	9	US-09-880-503-8
13	51.0	100.0	337	14	US-10-106-698-6266
14	51.0	100.0	337	15	US-10-264-049-2927
15	51.0	100.0	403	9	US-09-880-503-6

#### RESULT 1

US-09-880-503-1  
; Sequence 1, Application US/09880503  
; Patent No. US20020131964A1  
; GENERAL INFORMATION:  
; APPLICANT: CINES, Douglas B  
; APPLICANT: HIGAZI, Abd Al-Roof  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND

#### ALIGNMENTS

510	100.0	411	9	US-09-880-503-3	Sequence 3, Appli
510	100.0	431	9	US-09-264-488B-1	Sequence 1, Appli
510	100.0	431	12	US-10-411-037-34	Sequence 34, Appl
510	100.0	431	12	US-10-411-026-34	Sequence 34, Appl
510	100.0	431	13	US-10-076-421-2	Sequence 2, Appli
510	100.0	431	13	US-10-171-311-184	Sequence 184, App
510	100.0	431	14	US-10-193-656-4	Sequence 4, Appli
510	100.0	431	14	US-10-301-832-161	Sequence 161, App
510	100.0	431	14	US-10-131-985-21	Sequence 21, Appl
510	100.0	431	15	US-10-235-027-414	Sequence 414, App
510	100.0	431	15	US-10-295-027-1275	Sequence 1275, Ap
510	100.0	431	16	US-10-410-962-34	Sequence 34, Appl
510	100.0	431	16	US-10-411-049-34	Sequence 34, Appl
510	100.0	437	12	US-10-087-192-594	Sequence 594, App
508	99.6	431	14	US-10-247-671-149	Sequence 149, App
501	98.2	431	12	US-10-282-174-562	Sequence 562, App
500	98.0	411	15	US-10-407-821-2	Sequence 2, Appli
500	98.0	687	12	US-10-424-999-17	Sequence 17, Appl
500	98.0	687	12	US-10-425-000-37	Sequence 37, Appl
500	98.0	687	15	US-10-233-675A-17	Sequence 17, Appl
499	97.8	86	12	US-10-424-999-5	Sequence 5, Appli
499	97.8	86	12	US-10-424-999-62	Sequence 62, Appl
499	97.8	86	12	US-10-425-000-97	Sequence 97, Appl
499	97.8	86	15	US-10-233-675A-5	Sequence 5, Appli
499	97.8	86	15	US-10-233-675A-22	Sequence 22, Appl
499	97.8	87	12	US-10-424-999-9	Sequence 9, Appli
499	97.8	87	12	US-10-425-000-29	Sequence 29, Appl
499	97.8	87	15	US-10-233-675A-9	Sequence 9, Appli
499	97.8	91	12	US-10-424-999-37	Sequence 37, Appl
499	97.8	221	12	US-10-424-999-36	Sequence 36, Appl
499	97.8	445	15	US-10-360-101-266	Sequence 266, App
494	96.9	86	12	US-10-424-999-1	Sequence 1, Appli
494	96.9	86	15	US-10-233-675A-1	Sequence 1, Appli
494	96.9	87	12	US-10-424-999-10	Sequence 10, Appl
494	96.9	87	12	US-10-425-000-30	Sequence 30, Appl
494	96.9	87	15	US-10-233-675A-10	Sequence 10, Appl
494	96.9	91	12	US-10-424-999-35	Sequence 35, Appl
494	96.9	221	12	US-10-424-999-34	Sequence 34, Appl
494	96.9	322	12	US-10-424-999-20	Sequence 20, Appl
494	96.9	322	12	US-10-425-000-21	Sequence 21, Appl
494	96.9	322	12	US-10-425-000-40	Sequence 40, Appl
494	96.9	322	15	US-10-425-000-41	Sequence 41, Appl
494	96.9	322	15	US-10-233-675A-20	Sequence 20, Appl
494	96.9	322	15	US-10-233-675A-21	Sequence 21, Appl
494	96.9	672	12	US-10-424-999-15	Sequence 15, Appl
494	96.9	672	12	US-10-425-000-35	Sequence 35, Appl
494	96.9	672	15	US-10-233-675A-15	Sequence 15, Appl
494	96.9	674	12	US-10-424-999-14	Sequence 14, Appl
494	96.9	674	12	US-10-425-000-34	Sequence 34, Appl
494	96.9	688	12	US-10-233-675A-14	Sequence 14, Appl
494	96.9	688	12	US-10-424-999-18	Sequence 18, Appl
494	96.9	688	15	US-10-425-000-38	Sequence 38, Appl
494	96.9	688	15	US-10-233-675A-18	Sequence 18, Appl
494	96.9	689	12	US-10-424-999-13	Sequence 13, Appl
494	96.9	689	12	US-10-425-000-33	Sequence 33, Appl
494	96.9	689	15	US-10-233-675A-13	Sequence 13, Appl
494	96.7	650	15	US-10-401-077-1	Sequence 1, Appli
489	95.9	86	12	US-10-424-999-7	Sequence 7, Appli
489	95.9	86	15	US-10-233-675A-7	Sequence 7, Appli
483	94.7	86	15	US-10-233-675A-27	Sequence 27, Appl

X: Homo sapiens.  
C: EP405285-A.  
X: 02-JAN-1991.  
X: 18-JUN-1990; 90EP-00111471.  
X: 19-JUN-1989; 89JP-00156302.  
X: (KYOW ) KYOWA HAKKO KOGYO KK.  
X: Yasamura S, Nishi T, Ito S;  
D: WPI; 1991-008678/02.  
D: N-PSDB; AAQ10170.  
X: New plasminogen activator almost identical to natural pro:urokinase - is  
P: thrombin resistant and used for prophylaxis-treatment of cerebral  
P: thrombosis or myocardial infarction.  
X: Disclosure; Page 9; 84pp; English.  
X: UK-S3 is one example of a plasminogen activator which differs from  
C: natural human pro-urokinase at positions 153 and 155. (Leu substituted by  
C: Asn; Pro substituted by Thr, respectively). The derivative has decreased  
C: susceptibility to thrombin compared to natural type pro-UK and higher  
C: specific activity. See also AAQ10168 and AAQ10169  
X: Sequence 411 AA;  
S: Query Match 100.0%; Score 510; DB 2; Length 411;  
Best Local Similarity 100.0%; Pred. No. 9.2e-41;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Q: 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 60  
D: 48 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 107  
Q: 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
D: 108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135  
R: JLT 50  
A: 17956  
I: AAR47956 standard; protein; 411 AA.  
X: AAR47956;  
X: 10-FEB-1994 (first entry)  
X: PUK G16A G17A.  
D: Pre-urokinase; thrombolytic; blood; plasmid; mutant; PUK.  
X: Homo sapiens.  
F: Key Location/Qualifiers  
F: Domain 10..49  
F: /label= EGF  
F: Region 10..19  
F: /label= Loop\_1  
F: Region 20..31  
F: /label= Loop\_2  
F: Region 33..42  
F: /label= Loop\_3  
F: JP05192142-A.  
X: 03-AUG-1993.  
PF 20-JAN-1992; 92JP-00030178.  
XX  
PR 20-JAN-1992; 92JP-00030178.  
XX  
PA (GREG ) GREEN CROSS CORP.  
XX  
DR WPI; 1993-277461/35.  
XX  
PT Mutant human pre-urokinase - by replacing specified aminoacid(s) in N-  
PT terminal for providing longer half-life in blood and higher thrombolytic  
PT ability.  
XX  
XS Disclosure; Page 3; 26pp; Japanese.  
XX  
CC Human pre-urokinase can be modified by replacing: Ser, Asn, Pro, Gly, or  
CC Tyr in the region ranging from residue 10 to 60 in the N-terminal of the  
CC human PUK by Thr, Pro or Ala. These mutants have a longer half-life in  
CC the blood and a higher thrombolytic ability. For examples see (AAR47956-  
CC R47960)  
XX  
SQ Sequence 411 AA;  
Query Match 100.0%; Score 510; DB 2; Length 411;  
Best Local Similarity 100.0%; Pred. No. 9.2e-41;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 60  
DB 48 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 107  
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
DB 108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135  
Search completed: May 25, 2004, 14:52:51  
Job time : 22.518 secs

108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

ILT 47

3634

AAW13634 standard; protein; 411 AA.

AAW13634;

04-JUN-1997 (first entry)

Human native prourokinase.

Human; prourokinase; hPUK; variant; half-life; increase; EGF; epidermal growth factor domain; deletion; thrombolysis; fibrinolysis.

Homo sapiens

Key Location/Qualifiers

Domain 10..42  
/label= EGF domain  
/note= "in Claimed variants, at least part of the EGF domain is deleted (see comments)"

Region 10..19

Region 20..31

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42

Region 33..42



Sequence 411 AA;  
Query Match 100.0%; Score 510; DB 2; Length 411;  
Best Local Similarity 100.0%; Pred. No. 9.2e-41;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPDN 60  
48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPDN 107  
61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135  
RESULT 46  
AAR05117  
ID AAR05117 standard; protein; 411 AA.  
XX AAR05117;  
XX 25-MAR-2003 (revised)  
DT 04-OCT-1990 (first entry)  
XX  
DE UK-S3 as encoded by PUKS3.  
XX  
XW Urokinase; glycosylation.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 153  
FT /label= synthetic mutation  
FT /note= "old seq (Leu)"  
FT Misc-difference 155  
FT /label= synthetic mutation  
FT /note= "old seq (Pro)"  
PN BP370205-A.  
XX  
XX 30-MAY-1990.  
PD  
XX 28-SEP-1989; 89EP-00117981.  
PF  
XX 29-SEP-1988; 88JP-00245705.  
PR  
XX (KYOW ) KYOWA HAKKO KOGYO KK.  
PA  
XX Sasaki K, Nishi T, Yasumura S, Sato M, Itoh S;  
PI  
XX WPI; 1990-165029/22.  
DR  
XX N-PSDB; AAQ04486.  
XX  
PT Polypeptide(s) with added carbohydrate chains - formed by modification of  
PT aminoacid sequence, used to improve physio:chemical properties and/or  
PT activities.  
XX  
XX Disclosure; Page 7; 30pp; English.  
XX  
XX The polypeptide is a deriv. of mature urokinase, designated UK-S3 which  
XX has 2 amino acid subunits which result in an N-linked glycosylation site  
XX giving the new protein improved stability and activity. See also AAR05113  
XX -17. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-  
XX 2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)  
XX  
XX Sequence 411 AA;  
Query Match 100.0%; Score 510; DB 2; Length 411;  
Best Local Similarity 100.0%; Pred. No. 9.2e-41;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPDN 60  
48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPDN 107  
61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135  
RESULT 45  
AAR07903  
ID AAR07903 standard; protein; 411 AA.  
XX AAR07903;  
XX 21-FEB-1991 (first entry)  
DT  
XX Human pro-urokinase variant.  
XX  
XX Thrombin; fibrin; bleeding; pHR24.  
XX  
XX Homo sapiens  
XX  
XX Key Location/Qualifiers  
FH Domain 10..42  
FT /label= Epidermal growth factor (EGF) domain  
FT Region 10..19  
FT /label= First loop  
FT Region 20..31  
FT /label= Second loop  
FT Active-site 24..26  
FT /label= Modified site  
FT Region 33..42  
FT /label= Third loop  
PN EP398362-A.  
XX  
XX 22-NOV-1990.  
PD  
XX 18-MAY-1990; 90EP-00109473.  
PF  
XX 18-MAY-1989; 88JP-00126434.  
PR  
XX (GREG ) GREEN CROSS CORP.  
PA  
XX Tanabe T, Amatsuji Y, Kasai S, Hirose M, Morita M, Kawabe H;  
PI  
XX Arimura H;  
XX  
XX WPI; 1990-350147/47.  
DR  
XX N-PSDB; AAQ06134.  
XX  
PT Human pro-urokinase variant - produced by recombinant methods, showing  
PT increased half life in blood and high affinity for fibrin.  
XX  
XX Disclosure; Fig 1; 27pp; English.  
XX  
XX Modified pro-urokinase has a longer half-life in blood, and discloses  
XX thrombin without causing the spontaneous bleeding associated with  
XX urokinase. The modification puts an epidermal growth factor domain into  
XX the protein, containing the sequence Asn-X-Ser or Asn-X-Thr where X is  
XX any residue. Plasmid pHR24 is disclosed as containing the modified  
XX sequence  
XX  
XX Sequence 411 AA;  
Query Match 100.0%; Score 510; DB 2; Length 411;  
Best Local Similarity 100.0%; Pred. No. 9.2e-41;

Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPDN 60  
DB 48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPDN 107  
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
DB 108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135  
RESULT 46  
AAR05117  
ID AAR05117 standard; protein; 411 AA.  
XX AAR05117;  
XX 25-MAR-2003 (revised)  
DT 04-OCT-1990 (first entry)  
XX  
DE UK-S3 as encoded by PUKS3.  
XX  
XW Urokinase; glycosylation.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 153  
FT /label= synthetic mutation  
FT /note= "old seq (Leu)"  
FT Misc-difference 155  
FT /label= synthetic mutation  
FT /note= "old seq (Pro)"  
PN BP370205-A.  
XX  
XX 30-MAY-1990.  
PD  
XX 28-SEP-1989; 89EP-00117981.  
PF  
XX 29-SEP-1988; 88JP-00245705.  
PR  
XX (KYOW ) KYOWA HAKKO KOGYO KK.  
PA  
XX Sasaki K, Nishi T, Yasumura S, Sato M, Itoh S;  
PI  
XX WPI; 1990-165029/22.  
DR  
XX N-PSDB; AAQ04486.  
XX  
PT Polypeptide(s) with added carbohydrate chains - formed by modification of  
PT aminoacid sequence, used to improve physio:chemical properties and/or  
PT activities.  
XX  
XX Disclosure; Page 7; 30pp; English.  
XX  
XX The polypeptide is a deriv. of mature urokinase, designated UK-S3 which  
XX has 2 amino acid subunits which result in an N-linked glycosylation site  
XX giving the new protein improved stability and activity. See also AAR05113  
XX -17. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-  
XX 2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)  
XX  
XX Sequence 411 AA;  
Query Match 100.0%; Score 510; DB 2; Length 411;  
Best Local Similarity 100.0%; Pred. No. 9.2e-41;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPDN 60  
48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQGLGKHNYCRNPDN 107  
61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135  
RESULT 46  
AAR05117  
ID AAR05117 standard; protein; 411 AA.  
XX AAR05117;  
XX 25-MAR-2003 (revised)  
DT 04-OCT-1990 (first entry)  
XX  
DE UK-S3 as encoded by PUKS3.  
XX  
XW Urokinase; glycosylation.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 153  
FT /label= synthetic mutation  
FT /note= "old seq (Leu)"  
FT Misc-difference 155  
FT /label= synthetic mutation  
FT /note= "old seq (Pro)"  
PN BP370205-A.  
XX  
XX 30-MAY-1990.  
PD  
XX 28-SEP-1989; 89EP-00117981.  
PF  
XX 29-SEP-1988; 88JP-00245705.  
PR  
XX (KYOW ) KYOWA HAKKO KOGYO KK.  
PA  
XX Sasaki K, Nishi T, Yasumura S, Sato M, Itoh S;  
PI  
XX WPI; 1990-165029/22.  
DR  
XX N-PSDB; AAQ04486.  
XX  
PT Polypeptide(s) with added carbohydrate chains - formed by modification of  
PT aminoacid sequence, used to improve physio:chemical properties and/or  
PT activities.  
XX  
XX Disclosure; Page 7; 30pp; English.  
XX  
XX The polypeptide is a deriv. of mature urokinase, designated UK-S3 which  
XX has 2 amino acid subunits which result in an N-linked glycosylation site  
XX giving the new protein improved stability and activity. See also AAR05113  
XX -17. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-  
XX 2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)  
XX  
XX Sequence 411 AA;  
Query Match 100.0%; Score 510; DB 2; Length 411;  
Best Local Similarity 100.0%; Pred. No. 9.2e-41;

Claim 3; Fig 1; 11pp; English.

By forming a precursor-lipid composite, the half-life of this thrombolytic agent in the blood may be increased, exhibiting improved activity without abnormal acceleration of fibrinolytic activity. Compound is useful as a thrombolytic agent in treatment of cerebral thrombosis, myocardial infarction etc

Sequence 411 AA;

Query Match 100.0%; Score 510; DB 2; Length 411;  
Best Local Similarity 100.0%; Pred. No. 9.2e-41;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHFGYRGKASTDTMGPRCLPWNATVLOQTYHAHRSDALQGLGKKNYCRNPDN 60  
DB 48 KTCYEGNGHFGYRGKASTDTMGPRCLPWNATVLOQTYHAHRSDALQGLGKKNYCRNPDN 107  
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
DB 108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

LT 43  
M 7904  
II AAR07904 standard; protein; 411 AA.  
X AAR07904;

21-FEB-1991 (first entry)  
Human pro-urokinase variant.

Thrombin; fibrin; bleeding; pHR27.  
Homo sapiens.

Key Location/Qualifiers  
Domain 10..42  
Region 10..19  
Region 20..31  
Active-site 27..29  
Region 33..42  
EP398362-A.  
22-NOV-1990.  
18-MAY-1990; 90EP-00109473.  
18-MAY-1989; 89JP-00126434.  
(GRC) GREEN CROSS CORP.  
Tanabe T, Amatsuji Y, Kasai S, Hirose M, Morita M, Kawabe H;  
Arimura H;  
WPI; 1990-350147/47.  
N-PSDB; AAQ06135.

Human pro-urokinase variant - produced by recombinant methods, showing increased half life in blood and high affinity for fibrin.

Disclosure; Fig 1; 27pp; English.

Modified pro-urokinase has a longer half-life in blood, and dissolves thrombin without causing the spontaneous bleeding associated with urokinase. The modification puts an epidermal growth factor domain into

CC the protein, containing the sequence Asn-X-Ser or Asn-X-Thr where X is any residue. Plasmid pHR27 is disclosed as containing the modified sequence

SQ Sequence 411 AA;

Query Match 100.0%; Score 510; DB 2; Length 411;  
Best Local Similarity 100.0%; Pred. No. 9.2e-41;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFGYRGKASTDTMGPRCLPWNATVLOQTYHAHRSDALQGLGKKNYCRNPDN 60  
DB 48 KTCYEGNGHFGYRGKASTDTMGPRCLPWNATVLOQTYHAHRSDALQGLGKKNYCRNPDN 107

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
DB 108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

RESULT 44

AAR07902  
ID AAR07902 standard; protein; 411 AA.

AC AAR07902;

DT 21-FEB-1991 (first entry)

DE Human pro-urokinase variant.

XX Thrombin; fibrin; bleeding; pHR22.

OS Homo sapiens.

Key Location/Qualifiers  
Domain 10..42  
Region 10..19  
Region 20..31  
Active-site 22..24  
Region 33..42  
EP398362-A.  
22-NOV-1990.  
18-MAY-1990; 90EP-00109473.  
18-MAY-1989; 89JP-00126434.  
(GRC) GREEN CROSS CORP.  
Tanabe T, Amatsuji Y, Kasai S, Hirose M, Morita M, Kawabe H;  
Arimura H;  
WPI; 1990-350147/47.  
N-PSDB; AAQ06133.

Human pro-urokinase variant - produced by recombinant methods, showing increased half life in blood and high affinity for fibrin.

Disclosure; Fig 1; 27pp; English.

Modified pro-urokinase has a longer half-life in blood, and dissolves thrombin without causing the spontaneous bleeding associated with urokinase. The modification puts an epidermal growth factor domain into the protein, containing the sequence Asn-X-Ser or Asn-X-Thr where X is any residue. Plasmid pHR22 is disclosed as containing the modified sequence

XX 07-AUG-1990; 90JP-00207659.  
 XX 07-AUG-1990; 90JP-00207659.  
 XX (TOYJ) TOSOH CORP.  
 XX WPI; 1992-154820/19.  
 XX Vector contg. pro-urokinase encoding gene - includes mouse-IgG H chain E-mu enhancer, enhancer contg. SV40, early phase promoter and SV40 poly:adenylation site.  
 XX Claim 3; Page 1; 15pp; Japanese.  
 XX The sequence is that of mature prourokinase having the substitution mutations F157D and K15Q. The sequence per se is not given in the specification, so the known sequence of prourokinase was used to demonstrate the mutant protein. The gene encoding prourokinase has been over-expressed by recombinant DNA technology and provides a method for the rapid, low cost prodn. of prourokinase, which is a low mol. wt. plasminogen activator used for dissolving blood clots. See also AAR23794  
 XX Sequence 410 AA.  
 XX Query Match 100.0%; Score 510; DB 2; Length 410;  
 XX Best Local Similarity 100.0%; Pred. No. 9.2e-41;  
 XX Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKXNYCRNPDN 60  
 XX 47 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKXNYCRNPDN 106  
 XX 61 RRRPWCYVQVGLKPLVQECWVHDCADGK 88  
 XX 107 RRRPWCYVQVGLKPLVQECWVHDCADGK 134  
 XX  
 XX IIT 41  
 XX AAR50871 standard; protein; 411 AA.  
 XX AAP50871;  
 XX 30-NOV-1991 (first entry)  
 XX Sequence encoded by cDNA sequence for human urokinase zymogen (Japanese Patent Application No.37119/84).  
 XX Thrombolytic agent; plasminogen activator activity; fibrin affinity; enzyme.  
 XX Homo sapiens.  
 XX Key Location/Qualifiers  
 XX Disulfide-bond 50..131  
 XX Disulfide-bond 71..113  
 XX Disulfide-bond 102..126  
 XX Disulfide-bond 148..279  
 XX Cleavage-site 158..159  
 XX /note= "potential cleavage site which generates the two-chain form from the zymogen"  
 XX Disulfide-bond 189..205  
 XX Disulfide-bond 197..268  
 XX Disulfide-bond 293..362  
 XX Disulfide-bond 325..341  
 XX Disulfide-bond 352..380  
 XX EPI39447-A.  
 XX 02-MAY-1985.  
 XX 07-SEP-1984; 84EP-00306117.  
 XX  
 XX 13-SEP-1983; 83JP-00170354.  
 XX 17-OCT-1983; 83JP-00195051.  
 XX (GREG) GREEN CROSS CORP.  
 XX Kasai S, Arimura H, Mori K, Suyama T;  
 XX WPI; 1985-106530/18.  
 XX New urokinase zymogen - useful as thrombolytic agent.  
 XX Disclosure; Page 12; 30pp; English.  
 XX Zymogen AAP50871 is the inactive precursor form of human urokinase. Urokinase zymogen is cleaved into the two-chain form composed of characteristic urokinase H (molecular wt. of 30,000) and L (molecular wt. of 20,000) chains when treated with catalytic amounts of plasmin. The patentors claim a new urokinase zymogen which has mol. wt. ca. 50,000, a single chain molecular structure, and selective affinity for fibrin. It is a thrombolytic agent which manifests its plasminogen activator activity on cleavage by proteolytic enzymes (e.g. plasmin) and has higher affinity for fibrin than known forms of urokinase  
 XX Sequence 411 AA;  
 XX Query Match 100.0%; Score 510; DB 1; Length 411;  
 XX Best Local Similarity 100.0%; Pred. No. 9.2e-41;  
 XX Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKXNYCRNPDN 60  
 XX 48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKXNYCRNPDN 107  
 XX 61 RRRPWCYVQVGLKPLVQECWVHDCADGK 88  
 XX 108 RRRPWCYVQVGLKPLVQECWVHDCADGK 135  
 XX  
 XX RESULT 42  
 XX AAR06244  
 XX ID AAR06244 standard; protein; 411 AA.  
 XX AAR06244;  
 XX 07-DEC-1990 (first entry)  
 XX Urokinase precursor protein.  
 XX Urokinase precursor; fibrinolysis; thrombolytic; cerebral thrombosis; myocardial infarction.  
 XX Homo sapiens.  
 XX EP380334-A.  
 XX 01-AUG-1990.  
 XX 25-JAN-1990; 90EP-00300772.  
 XX 27-JAN-1989; 89JP-00016406.  
 XX 17-MAY-1989; 89JP-00121405.  
 XX (GREG) GREEN CROSS CORP.  
 XX Matsuda H, Ueda Y, Tamanouchi K;  
 XX WPI; 1990-233117/31.  
 XX Urokinase precursor-lipid composite - used as thrombolytic agent, having prolonged half-life in the blood, enhanced bioavailability and improved activity.

LT 38  
547  
AAE16547 standard; protein; 403 AA.  
AAE16547;  
09-APR-2002 (first entry)  
Human urokinase-type plasminogen activator scupA delta136-143 mutant.  
Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;  
stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;  
microvascular occlusion; angiodysplasia; pulmonary fibrosis; asthma;  
tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;  
clotting disorder; uterine contraction disorder; respiratory disease;  
male impotence; adult respiratory distress syndrome; scupA delta136-143;  
single chain urokinase; mutant; mutein.  
Homo sapiens.  
Synthetic.  
W0200197752-A2.  
27-DEC-2001.  
13-JUN-2001; 2001WO-018976.  
20-JUN-2000; 2000US-0212874P.  
(UYPE-) UNIV PENNSYLVANIA.  
Cines DB, Higazi AA;  
WPI; 2002-122240/16.  
N-PSDB; AAD27080.  
Composition for modulating muscle cell and tissue contractility for  
treating atherosclerosis, asthma, hypertension, glaucoma, impotence,  
comprising domains from urokinase-type plasminogen activator.  
Claim 22; Fig 1F; 117pp; English.  
The invention relates to a composition comprising one or more domains of  
urokinase-type plasminogen activator (uPA). The composition is used to  
modulate the contractility and angiogenic activity of a mammalian muscle,  
endothelial cell or tissue. The composition is used for treating stroke,  
hypotension, hypertension, atherosclerosis, heart attack, microvascular  
occlusions, thrombotic microangiopathies, surgically induced thrombotic  
disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell  
invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,  
diabetic retinopathy, wound healing, clotting disorder, uterine  
contraction disorder, male impotence, respiratory disease or condition  
such as asthma, adult respiratory distress syndrome, primary pulmonary  
hypertension, microvascular thrombotic occlusion, and a disorder  
associated with chronic intrapulmonary fibrin formation. The present  
sequence is human urokinase-type plasminogen activator (uPA) single chain  
urokinase (scuPA) deletion mutant designated as scuPA delta136-143  
Sequence 403 AA;  
Query Match 100.0%; Score 510; DB 5; Length 403;  
Best Local Similarity 100.0%; Pred. No. 9e-41;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 KTCYEGNGHFGKASTDTMGRCPLPWSATVLCQTYHAHRSDALQLGLGKHNYCRNPDN 60  
48 KTCYEGNGHFGKASTDTMGRCPLPWSATVLCQTYHAHRSDALQLGLGKHNYCRNPDN 107  
61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

RESULT 39  
AAR23794  
ID AAR23794 standard; protein; 410 AA.  
XX  
AC AAR23794;  
XX  
DT 03-NOV-1992 (first entry)  
XX  
DE Prourokinase mutant F157D.  
XX  
KW Substitution; animal; plasminogen activator; blood; clot.  
XX  
OS Homo sapiens.  
XX  
PN JP04091792-A.  
XX  
PD 25-MAR-1992.  
XX  
PF 07-AUG-1990; 90JP-00207659.  
XX  
PR 07-AUG-1990; 90JP-00207659.  
XX  
PA (TOYJ ) TOSOH CORP.  
XX  
DR WPI; 1992-154820/19.  
XX  
PT Vector contg. pro-urokinase encoding gene - includes mouse-IgG H chain E-  
mu enhancer, enhancer contg. SV40, early phase promoter and SV40  
poly:adenylation site.  
XX  
PS Claim 3; Page 1; 15pp; Japanese.  
XX  
CC The sequence is that of mature prourokinase having the substitution  
mutation F157D. The sequence per se is not given in the specification, so  
the known sequence of prourokinase was used to demonstrate the mutant  
protein. The gene encoding prourokinase has been over-expressed by  
CC recombinant DNA technology and provides a method for the rapid, low cost  
CC prodn. of prourokinase, which is a low mol. wt. plasminogen activator  
CC used for dissolving blood clots. See also AAR23795  
XX  
SQ Sequence 410 AA;  
Query Match 100.0%; Score 510; DB 2; Length 410;  
Best Local Similarity 100.0%; Pred. No. 9.2e-41;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHFGKASTDTMGRCPLPWSATVLCQTYHAHRSDALQLGLGKHNYCRNPDN 60  
Db 47 KTCYEGNGHFGKASTDTMGRCPLPWSATVLCQTYHAHRSDALQLGLGKHNYCRNPDN 106  
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
Db 107 RRRPWCYVQVGLKPLVQECMVHDCADGK 134  
RESULT 40  
AAR23795  
ID AAR23795 standard; protein; 410 AA.  
XX  
AC AAR23795;  
XX  
DT 03-NOV-1992 (first entry)  
XX  
DE Prourokinase double mutant F157D, K135Q.  
XX  
KW Substitution; animal; plasminogen activator; blood; clot.  
XX  
OS Homo sapiens.  
XX  
PN JP04091792-A.  
XX  
PD 25-MAR-1992.

D' 25-MAR-2003 (revised)  
 D' 17-AUG-1995 (first entry)  
 X' Bifunctional urokinase variant M15.  
 X' fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
 K' urokinase; variant; mutein.  
 X' Synthetic.  
 O' Key  
 F' Region  
 F' Location/Qualifiers  
 F' 1..365  
 F' /label= M4  
 F' /note= "unglycosylated prourokinase (Ser47-Leu411)"  
 F' Disulfide-bond 4..85  
 F' Disulfide-bond 25..67  
 F' Disulfide-bond 56..80  
 F' Disulfide-bond 102..233  
 F' Disulfide-bond 143..159  
 F' Disulfide-bond 151..222  
 F' Disulfide-bond 247..316  
 F' Disulfide-bond 279..295  
 F' Disulfide-bond 306..334  
 F' Disulfide-bond 366..378  
 F' Region  
 F' /label= X1  
 F' 379..397  
 F' /label= Y1  
 X' DE4323754-C1  
 X' 01-DEC-1994  
 X' 15-JUL-1993; 93DE-04323754.  
 X' 15-JUL-1993; 93DE-04323754.  
 X' (CHEF ) GRUENENTHAL GMBH.  
 X' Steffens GJ, Wnendt S, Schneider J, Heinzl-Wieland R;  
 X' Saunders DJ;  
 X' WPI; 1995-015191/03.  
 X' New bifunctional urokinase derivs and related plasmids - with improved  
 F' fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
 F' cerebral infarct, pulmonary embolism, etc.  
 X' Example 1; Page 10 and Fig 1; 34pp; German.  
 X' Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
 C' are claimed (see features table). Sequences AAR6244-R66266 are specific  
 C' examples of such derivs, which have both improved fibrinolytic and  
 C' thrombin-inhibiting activities, compared to known plasminogen activators  
 C' or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
 C' e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
 C' and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
 C' to correct PN field.)  
 X' Sequence 397 AA;  
 S' Query Match 100.0%; Score 510; DB 2; Length 397;  
 S' Best Local Similarity 100.0%; Pred. No. 8.9e-41;  
 S' Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Q 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRS DALQLGLGKHNYCRNPDN 60  
 D 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRS DALQLGLGKHNYCRNPDN 61  
 Q 61 RRRPWCYVQVGLKPLVQECMWHDCADGK 88  
 D 62 RRRPWCYVQVGLKPLVQECMWHDCADGK 89

RESULT 37  
 AAW13637  
 ID AAW13637 standard; protein; 401 AA.  
 XX  
 AC AAW13637;  
 XX  
 DT 04-JUN-1997 (first entry)  
 XX  
 DE Human prourokinase variant lacking EGF domain loop 3.  
 DE  
 KW Human; prourokinase; hPUK; variant; half-life; increase; EGF;  
 KW epidermal growth factor domain; deletion; thrombolysis; fibrinolysis.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 FH Key  
 FH Region  
 FT 1..32 Location/Qualifiers  
 FT /note= "residues 1-32 of native hPUK"  
 FT Region 33..401  
 FT /note= "residues 43-411 of native hPUK"  
 FT Misc-difference 167  
 FT /note= "corresponds to TAC codon"  
 XX  
 PN EP398361-A.  
 XX  
 PD 22-NOV-1990.  
 XX  
 PF 18-MAY-1990; 90EP-00109472.  
 XX  
 PR 18-MAY-1989; 89JP-00126433.  
 XX  
 PR 22-FEB-1990; 90JP-00042020.  
 XX  
 PA (GREC ) GREEN CROSS CORP.  
 XX  
 PI Tanabe T, Amatsuji Y, Kasai S, Hirose M, Morita M, Kawabe H;  
 PI Airmura H;  
 XX  
 DR WPI; 1990-350146/47.  
 DR N-PSDB; AAT61674.  
 XX  
 PT Human pro-urokinase variants - deficient in loop regions of epidermal  
 PT growth factor, showing long blood half-life, as fibrinolytic agent.  
 XX  
 PS Claim 11; Page; 22pp; English.  
 XX  
 CC New variants of human prourokinase (hPUK) comprise a hPUK deficient in  
 CC (i) at least part of the first loop region of the epidermal growth factor  
 CC (EGF) domain; (ii) at least part of the first loop and at least part of  
 CC the second loop; or (iii) at least part of the third loop. The hPUK  
 CC variants show an increased blood half-life comparable to that of the  
 CC whole EGF domain-deficient hPUK variant and urokinase while retaining the  
 CC same properties as those of hPUK. They have potent thrombolytic activity  
 CC and very little tendency to cause spontaneous bleeding. The present  
 CC sequence represents a specific variant of hPUK which lacks EGF domain  
 CC loop 3; the sequence does not appear in the specification and has been  
 CC created using the wild-type hPUK sequence and the junction sequence after  
 CC deletion, both of which are given (in Fig 1 and on page 8, respectively)  
 XX  
 SQ Sequence 401 AA;  
 Query Match 100.0%; Score 510; DB 2; Length 401;  
 Best Local Similarity 100.0%; Pred. No. 9e-41;  
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRS DALQLGLGKHNYCRNPDN 60  
 DB 38 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRS DALQLGLGKHNYCRNPDN 97  
 QY 61 RRRPWCYVQVGLKPLVQECMWHDCADGK 88  
 DB 98 RRRPWCYVQVGLKPLVQECMWHDCADGK 125

```

F) /label= M4
F) /note= "unglycosylated prourokinase(Ser47-Leu411)"
F) 4. .85
F) Disulfide-bond 102. .233
F) Disulfide-bond 143. .159
F) Disulfide-bond 151. .222
F) Disulfide-bond 247. .316
F) Disulfide-bond 279. .295
F) Disulfide-bond 306. .334
F) Disulfide-bond 366. .371
F) Region /label= X1
F) Region /label= Y1
F) DE4323754-C1.
F) 01-DEC-1994.
F) 15-JUL-1993; 93DE-04323754.
F) 15-JUL-1993; 93DE-04323754.
F) (CHEF ) GRUENENTHAL GMBH.
F) Steffens GJ, Wrendt S, Schneider J, Heinzel-Wieland R;
F) Saunders DJ;
F) WPI; 1995-015191/03.
F) New bifunctional urokinase derivs and related plasmids - with improved
F) fibrinolytic and thrombin inhibiting activities, for treating cardiac and
F) cerebral infarct, pulmonary embolism, etc.
F) Example 1; Page 11 and Fig 1; 34pp; German.
F) Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1
F) are claimed (see features table). Sequences AAR66244-R66266 are specific
F) examples of such derivs. which have both improved fibrinolytic and
F) thrombin-inhibiting activities, compared to known plasminogen activators
F) or thrombin inhibitors. The proteins are useful as thrombolytic agents,
F) e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac
F) and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003
F) to correct PN field.)
F) Sequence 395 AA;
F) Query Match 100.0%; Score 510; DB 2; Length 395;
F) at Local Similarity 100.0%; Pred. No. 8.9e-41;
F) tches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
F) 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOYTHAHRSDALQLGLGKHNCRPN 60
F) 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOYTHAHRSDALQLGLGKHNCRPN 61
F) 61 RRRPWCYVQGLKPLVQECMVHDCADGK 88
F) 62 RRRPWCYVQGLKPLVQECMVHDCADGK 89
F) LT 35
F) 6246
F) AAR66246 standard; protein; 396 AA.
F) AAR66246;
F) 25-MAR-2003 (revised)
F) 17-AUG-1995 (first entry)
F) Bifunctional urokinase variant M13.
F) fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;

```

```

KX urokinase; variant; mutein.
XX Synthetic.
XX Key Location/Qualifiers
XX Region /label= M4
XX /note= "unglycosylated prourokinase(Ser47-Leu411)"
XX Disulfide-bond 102. .233
XX Disulfide-bond 143. .159
XX Disulfide-bond 151. .222
XX Disulfide-bond 247. .316
XX Disulfide-bond 279. .295
XX Disulfide-bond 306. .334
XX Disulfide-bond 366. .371
XX Region /label= X1
XX Region /label= Y1
XX DE4323754-C1.
XX 01-DEC-1994.
XX 15-JUL-1993; 93DE-04323754.
XX 15-JUL-1993; 93DE-04323754.
XX (CHEF ) GRUENENTHAL GMBH.
XX Steffens GJ, Wrendt S, Schneider J, Heinzel-Wieland R;
XX Saunders DJ;
XX WPI; 1995-015191/03.
XX New bifunctional urokinase derivs and related plasmids - with improved
XX fibrinolytic and thrombin inhibiting activities, for treating cardiac and
XX cerebral infarct, pulmonary embolism, etc.
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1
XX are claimed (see features table). Sequences AAR66244-R66266 are specific
XX examples of such derivs. which have both improved fibrinolytic and
XX thrombin-inhibiting activities, compared to known plasminogen activators
XX or thrombin inhibitors. The proteins are useful as thrombolytic agents,
XX e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac
XX and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003
XX to correct PN field.)
XX Sequence 396 AA;
XX Query Match 100.0%; Score 510; DB 2; Length 396;
XX Best Local Similarity 100.0%; Pred. No. 8.9e-41;
XX Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOYTHAHRSDALQLGLGKHNCRPN 60
DB 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOYTHAHRSDALQLGLGKHNCRPN 61
QY 61 RRRPWCYVQGLKPLVQECMVHDCADGK 88
DB 62 RRRPWCYVQGLKPLVQECMVHDCADGK 89
RESULT 36
AAR66248
ID AAR66248 standard; protein; 397 AA.
XX AAR66248;
XX AAR66248;

```

X AAR47902;  
 X 13-JUL-1994 (first entry)  
 X Pro-urokinase derivative.  
 X Pro-urokinase; half-life; thrombolytic; thrombosis; fibrinolytic; factor.  
 X Homo sapiens  
 X JP05336965-A.  
 X 21-DEC-1993.  
 X 17-OCT-1991; 91JP-00269615.  
 X 17-OCT-1991; 91JP-00269615.  
 X (KYOWA) KYOWA HAKKO KOGYO KK.  
 X WPI; 1994-030907/04.  
 X N-PSDB; AAR55771.  
 X Novel human pro-urokinase derivs. having long half-life - with high  
 X thrombolytic activity, useful for treatment of thrombosis.  
 X Disclosure; Page 14; 29pp; Japanese.  
 X Sequences (AAR55771-72) are pro-urokinase derivatives. The products have  
 X an inserted sugar moiety having an amino acid substituted, depleted or  
 X inserted variant around the thrombin cleavage site. They also have a long  
 X half-life allowing them to be used in the treatment of thrombosis  
 X  
 X Sequence 395 AA;  
 X Query Match 100.0%; Score 510; DB 2; Length 395;  
 X Best Local Similarity 100.0%; Pred. No. 8.9e-41;  
 X Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 X  
 X 1 KTCYEGNGHFRGKASTDTMGRCPLPWSATVLOQTYHAHRSALQGLGKHNYCRNPDN 60  
 X 32 KTCYEGNGHFRGKASTDTMGRCPLPWSATVLOQTYHAHRSALQGLGKHNYCRNPDN 91  
 X 61 RRRPWCYVQVGLKPLVQECWVHDCADGK 88  
 X 92 RRRPWCYVQVGLKPLVQECWVHDCADGK 119  
 X  
 X JLT 33  
 X 6265  
 X AAR66265 standard; protein; 395 AA.  
 X AAR66265;  
 X 25-MAR-2003 (revised)  
 X 17-AUG-1995 (first entry)  
 X Bifunctional urokinase variant M32.  
 X fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
 X urokinase; variant; mutein.  
 X Synthetic.  
 X  
 X Key Location/Qualifiers  
 X Region 1..365  
 X /label= M4  
 X /note= "unglycosylated prourokinase (Ser47-Leu411)"  
 X Disulfide-bond 4..85  
 X Disulfide-bond 25..87  
 X Disulfide-bond 56..80  
 X Disulfide-bond 102..233

FT Disulfide-bond 143..159  
 FT Disulfide-bond 151..222  
 FT Disulfide-bond 247..316  
 FT Disulfide-bond 279..295  
 FT Disulfide-bond 306..334  
 FT Region 366..371  
 FT /label= X1  
 FT Region 372..395  
 FT /label= Y1  
 XX  
 XX DE4323754-Cl.  
 XX  
 XX 01-DEC-1994.  
 XX  
 XX 15-JUL-1993; 93DE-04323754.  
 XX  
 XX 15-JUL-1993; 93DE-04323754.  
 XX  
 XX (CHEF) GRUENENTHAL GMBH.  
 XX  
 XX Steffens GJ, Whendt S, Schneider J, Heinzel-Wieland R;  
 XX Saunders DJ;  
 XX WPI; 1995-015191/03.  
 XX  
 XX New bifunctional urokinase derivs and related plasmids - with improved  
 XX fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
 XX cerebral infarct, pulmonary embolism, etc.  
 XX  
 XX Example 1; Page 11 and Fig 1; 34pp; German.  
 XX  
 XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
 XX are claimed (see features table). Sequences AAR66244-R66266 are specific  
 XX examples of such derivs. which have both improved fibrinolytic and  
 XX thrombin-inhibiting activities, compared to known plasminogen activators  
 XX or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
 XX e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
 XX and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
 XX to correct PN field.)  
 XX  
 XX Sequence 395 AA;  
 XX Query Match 100.0%; Score 510; DB 2; Length 395;  
 XX Best Local Similarity 100.0%; Pred. No. 8.9e-41;  
 XX Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 XX 1 KTCYEGNGHFRGKASTDTMGRCPLPWSATVLOQTYHAHRSALQGLGKHNYCRNPDN 60  
 XX 2 KTCYEGNGHFRGKASTDTMGRCPLPWSATVLOQTYHAHRSALQGLGKHNYCRNPDN 61  
 XX 61 RRRPWCYVQVGLKPLVQECWVHDCADGK 88  
 XX 62 RRRPWCYVQVGLKPLVQECWVHDCADGK 89  
 XX  
 XX RESULT 34  
 XX AAR66262  
 XX ID AAR66262 standard; protein; 395 AA.  
 XX  
 XX AAR66262;  
 XX 25-MAR-2003 (revised)  
 XX 17-AUG-1995 (first entry)  
 XX  
 XX Bifunctional urokinase variant M29.  
 XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
 XX urokinase; variant; mutein.  
 XX Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 XX Region 1..365

Query Match 100.0%; Score 510; DB 2; Length 393;  
 Best Local Similarity 100.0%; Pred. No. 8.8e-41;  
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 60  
 D 3 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 62

Q 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
 D 63 RRRPWCYVQVGLKPLVQECMVHDCADGK 90

R1 LT 30  
 A1 9596  
 I1 AAR99596 standard; protein; 393 AA.  
 X1 AAR99596;  
 D1 05-DEC-1996 (first entry)  
 X1 Chimeric protein M37 encoded by pSE9.  
 D1 Thrombin; inhibition; thrombus; thrombolysis; chimeric protein;  
 K1 plasminogen activating sequence; fibrinolysis; infarction;  
 K1 angina pectoris; deep vein thrombosis.  
 X1 Synthetic.  
 O1 EP714982-A2.  
 X1  
 F1 05-JUN-1996.  
 P1 16-NOV-1995; 95EP-00118050.  
 X1  
 P1 30-NOV-1994; 94DE-04442665.  
 X1  
 P1 (CHEF ) GRUENTHAL GMBH.  
 X1  
 P1 Wnendt S, Steffens GJ, Janocha E, Heinzel-Wieland R;  
 X1 WPI; 1996-269715/28.  
 D1 Chimeric protein contg. plasminogen activating sequence and thrombin-  
 P1 inhibiting sequence - useful as thrombus-specific thrombolytic agent with  
 P1 rapid action.  
 X1  
 P1 Example 1; Page 19-20; 37pp; German.  
 X1  
 C1 Example 1 describes the prodn. of plasmids pSE1 and pSE9 contg. a DNA  
 C1 encoding a chimeric protein with fibrinolytic and thrombin-inhibiting  
 C1 properties. pSE1 encodes the protein given in AAR99597 and pSE9 encodes  
 C1 the protein given in AAR99596  
 X1  
 S1 Sequence 393 AA;

Query Match 100.0%; Score 510; DB 2; Length 393;  
 Best Local Similarity 100.0%; Pred. No. 8.8e-41;  
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 60  
 D 3 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 62

Q 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
 D 63 RRRPWCYVQVGLKPLVQECMVHDCADGK 90

R1 LT 31  
 A1 9885  
 I1 AAR99885 standard; peptide; 393 AA.  
 X1

AAR99885;  
 27-JAN-1997 (first entry)  
 M36: fibrinolytic and anticoagulant activity contg. protein.  
 Thrombin; factor Xa; anticoagulant; thrombolytic; plasminogen; urokinase;  
 activator; streptokinase; staphylokinase; APSAC;  
 anisolated plasminogen streptokinase activator complex; hirudin;  
 hirullin; antistatin; pWLT27; pWS1; pSE8; pHW56.  
 Synthetic.  
 EP712934-A2.  
 22-MAY-1996.  
 03-NOV-1995; 95EP-00117316.  
 17-NOV-1994; 94DE-04440892.  
 (CHEF ) GRUENTHAL GMBH.  
 Wnendt S, Heinzel-Wieland R, Steffens GJ;  
 WPI; 1996-240720/25.  
 Proteins with fibrinolytic and anticoagulant activity - useful as  
 thrombolytic agents.  
 Disclosure, Fig 18; 59pp; German.  
 New peptides (I) with fibrinolytic and anticoagulant activity comprise a  
 plasminogen-activating amino acid sequence (A) fused at the N- and/or C-  
 terminus to a thrombin and/or factor Xa inhibiting amino acid sequence  
 (B). Excluded from the claims are (I) where (A) is Ser47 to Leu411 of  
 unglycosylated urokinase linked at the C-terminus to sequences (i) to  
 (iii): Ti-RP-12-GGGGNGDFEIPPEYL-P3 (i) Ti-RPFLRPNDKIFRWFDEEKNE (ii)  
 Ti-RPSEFEFEIDEKEK (iii) Where Ti= P or V; T2 = L or a bond; T3= Q or  
 OH. (A) is pref. (pro)urokinase; tissue plasminogen activator (tPA), bat-  
 PA (all opt. modified by deletion, substitution, insertion and/or addn.);  
 streptokinase; staphylokinase; and/or APSAC (anisolated plasminogen  
 streptokinase activator complex), esp. prourokinase (411 amino acids) or  
 its Ser47 to Leu411 or Ser138 to Leu411 fragments, or t-PA (527 amino  
 acids) or its Ser89Arg to 527Pro or 174Ser to 527Pro fragments. (B) has  
 hirudin or hirullin activities; or is derived from the human thrombin  
 receptor, antistatin and/or the tick anticoagulant peptide. Most pref.  
 are the 65 amino acid hirudin sequence or one of the six sequences given  
 in AAR99879 to AAR99884. Plasmids pWLT27 (WS1), pWS1 (WS112), pSE8 (M36)  
 and pHW56 (M43) contain the sequences encoding AAR99885 to AAR99888,  
 respectively. The products were tested in human citrated plasma (5 microg  
 in 200 microl 1:10 diluted plasma). The thrombin time was then 1.2, 3,  
 2.8 and 1.2 times greater, respectively, than in the absence of the  
 product  
 Sequence 393 AA;

Query Match 100.0%; Score 510; DB 2; Length 393;  
 Best Local Similarity 100.0%; Pred. No. 8.8e-41;  
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 60  
 D 3 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNYCRNPDN 62

Q 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
 D 63 RRRPWCYVQVGLKPLVQECMVHDCADGK 90

RESULT 32  
 AAR47902  
 ID AAR47902 standard; protein; 395 AA.



C Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
C are claimed (see features table). Sequences AAR66244-R66266 are specific  
C examples of such derivs. which have both improved fibrinolytic and  
C thrombin-inhibiting activities, compared to known plasminogen activators  
C or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
C e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
C and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
C to correct PN field.)  
C  
C Sequence 393 AA;  
C  
C Query Match 100.0%; Score 510; DB 2; Length 393;  
C Best Local Similarity 100.0%; Pred. No. 8.8e-41;  
C Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
C  
C 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLTQTYHAHRSDALQLGLGKHNYCRNPDN 60  
C 2 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLTQTYHAHRSDALQLGLGKHNYCRNPDN 61  
C  
C 61 RRRPWCYVQVGLKPLVQECMWHDCADGK 88  
C 62 RRRPWCYVQVGLKPLVQECMWHDCADGK 89  
C  
C  
C ILT 28  
C AAR66250 standard; protein; 393 AA.  
C AAR66250;  
C  
C 25-MAR-2003 (revised)  
C 17-AUG-1995 (first entry)  
C  
C Bifunctional urokinase variant M17.  
C  
C fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
C urokinase; variant; mutein.  
C  
C Synthetic.  
C  
C Key Location/Qualifiers  
C Region 1..365 /label= M4  
C Disulfide-bond 4..85 /notes= "unglycosylated prourokinase(Ser47-Leu411)"  
C Disulfide-bond 25..67  
C Disulfide-bond 56..80  
C Disulfide-bond 102..233  
C Disulfide-bond 143..159  
C Disulfide-bond 151..222  
C Disulfide-bond 247..316  
C Disulfide-bond 279..295  
C Disulfide-bond 306..334  
C Disulfide-bond 366..372  
C Region /label= X1  
C Region 373..393 /label= Y1  
C  
C DB4323754-C1.  
C  
C 01-DEC-1994.  
C  
C 15-JUL-1993; 93DE-04323754.  
C  
C 15-JUL-1993; 93DE-04323754.  
C  
C (CHEF ) GRUENENTHAL GMBH.  
C  
C Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
C Saunders DJ;  
C WPI; 1995-015191/03.  
C  
C X

PT New bifunctional urokinase derivs and related plasmids - with improved  
PT fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
PT cerebral infarct, pulmonary embolism, etc.  
XX Example 1; Page 10 and Fig 1; 34pp; German.  
XX  
CC Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
CC are claimed (see features table). Sequences AAR66244-R66266 are specific  
CC examples of such derivs. which have both improved fibrinolytic and  
CC thrombin-inhibiting activities, compared to known plasminogen activators  
CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
CC to correct PN field.)  
XX  
SQ Sequence 393 AA;  
C  
C Query Match 100.0%; Score 510; DB 2; Length 393;  
C Best Local Similarity 100.0%; Pred. No. 8.8e-41;  
C Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
C  
C 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLTQTYHAHRSDALQLGLGKHNYCRNPDN 60  
C 2 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLTQTYHAHRSDALQLGLGKHNYCRNPDN 61  
C  
C 61 RRRPWCYVQVGLKPLVQECMWHDCADGK 88  
C 62 RRRPWCYVQVGLKPLVQECMWHDCADGK 89  
C  
C  
C RESULT 29  
C AAR99597  
C ID AAR99597 standard; protein; 393 AA.  
C AC AAR99597;  
C XX  
C DT 05-DEC-1996 (first entry)  
C XX  
C DE Chimeric protein M38 encoded by pSEI.  
C XX  
C KW Thrombin; inhibition; thrombus; thrombolysis; chimeric protein;  
C KW plasminogen activating sequence; fibrinolysis; infarction;  
C KW angina pectoris; deep vein thrombosis.  
C XX  
C OS Synthetic.  
C XX  
C PN EP714982-A2.  
C XX  
C PD 05-JUN-1996.  
C XX  
C PF 16-NOV-1995; 95EP-00118050.  
C XX  
C PR 30-NOV-1994; 94DE-04442665.  
C XX  
C PA (CHEF ) GRUENENTHAL GMBH.  
C XX  
C PI Wnendt S, Steffens GJ, Janocha E, Heinzel-Wieland R;  
C XX WPI; 1996-269715/28.  
C XX  
C PT Chimeric protein contg. plasminogen activating sequence and thrombin-  
C PT inhibiting sequence - useful as thrombus-specific thrombolytic agent with  
C PT rapid action.  
C XX  
C PS Example 1; Page 21-22; 37pp; German.  
C XX  
CC Example 1 describes the prodn. of plasmids pSEI and pSE9 contg. a DNA  
CC encoding a chimeric protein with fibrinolytic and thrombin-inhibiting  
CC properties. pSEI encodes the protein given in AAR99597 and pSE9 encodes  
CC the protein given in AAR99596  
C  
C Sequence 393 AA;  
C  
C X

Best Local Similarity 100.0%; Pred. No. 8.8e-41; Indels 0; Gaps 0;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWSATVLQOTYHAHRSALQLGLGKHNCRNPDN 60  
D 2 KTCYEGNGHFYRGKASTDTMGPRCLPWSATVLQOTYHAHRSALQLGLGKHNCRNPDN 61  
QY 61 RRRPWCYVQVGLKPLVQECWVHDCADGK 88  
D 62 RRRPWCYVQVGLKPLVQECWVHDCADGK 89  
R 26  
A 6252  
II AAR66252 standard; protein; 393 AA.  
A AAR66252;  
X 25-MAR-2003 (revised)  
D 17-AUG-1995 (first entry)  
X Bifunctional urokinase variant M19.  
X fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
K urokinase; variant; mutein.  
X Synthetic.  
X Key Location/Qualifiers  
F 1. 365  
F Region /label= M4  
F /note= "unglycosylated prourokinase (Ser47-Leu411)"  
F Disulfide-bond 4. .85  
F Disulfide-bond 25. .67  
F Disulfide-bond 56. .80  
F Disulfide-bond 102. .233  
F Disulfide-bond 143. .159  
F Disulfide-bond 151. .222  
F Disulfide-bond 247. .316  
F Disulfide-bond 279. .295  
F Disulfide-bond 306. .334  
F Disulfide-bond 366. .372  
F Region /label= X1  
F Region 373. .393  
F /label= Y1  
X DE4323754-Cl.  
X 01-DEC-1994.  
X 15-JUL-1993; 93DE-04323754.  
X 15-JUL-1993; 93DE-04323754.  
X (CHEF ) GRUENENTHAL GMBH.  
X Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
X Saunders DJ;  
X WPI; 1995-015191/03.  
X New bifunctional urokinase derivs and related plasmids - with improved  
X fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
X cerebral infarct, pulmonary embolism, etc.  
X Example 1; Page 10 and Fig 1; 34pp; German.  
X Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
X are claimed (see features table). Sequences AAR66244-R66265 are specific  
X examples of such derivs. which have both improved fibrinolytic and  
X thrombin-inhibiting activities, compared to known plasminogen activators  
X or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
X e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac

CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
CC to correct PN field.)  
XQ Sequence 393 AA;  
Query Match 100.0%; Score 510; DB 2; Length 393;  
Best Local Similarity 100.0%; Pred. No. 8.8e-41; Indels 0; Gaps 0;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWSATVLQOTYHAHRSALQLGLGKHNCRNPDN 60  
D 2 KTCYEGNGHFYRGKASTDTMGPRCLPWSATVLQOTYHAHRSALQLGLGKHNCRNPDN 61  
QY 61 RRRPWCYVQVGLKPLVQECWVHDCADGK 88  
D 62 RRRPWCYVQVGLKPLVQECWVHDCADGK 89  
RESULT 27  
AAR66244  
ID AAR66244 standard; protein; 393 AA.  
X AAR66244;  
X 25-MAR-2003 (revised)  
D 22-AUG-1995 (first entry)  
X Bifunctional urokinase variant M11.  
X fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
K urokinase; variant; mutein.  
X Synthetic.  
X Key Location/Qualifiers  
F 1. 365  
F Region /label= M4  
F /note= "unglycosylated prourokinase (Ser47-Leu411)"  
F Disulfide-bond 4. .85  
F Disulfide-bond 25. .67  
F Disulfide-bond 56. .80  
F Disulfide-bond 102. .233  
F Disulfide-bond 143. .159  
F Disulfide-bond 151. .222  
F Disulfide-bond 247. .316  
F Disulfide-bond 279. .295  
F Disulfide-bond 306. .334  
F Disulfide-bond 366. .374  
F Region /label= X1  
F Region 375. .393  
F /label= Y1  
X DE4323754-Cl.  
X 01-DEC-1994.  
X 15-JUL-1993; 93DE-04323754.  
X 15-JUL-1993; 93DE-04323754.  
X (CHEF ) GRUENENTHAL GMBH.  
X Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
X Saunders DJ;  
X WPI; 1995-015191/03.  
X New bifunctional urokinase derivs and related plasmids - with improved  
X fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
X cerebral infarct, pulmonary embolism, etc.  
X Example 1; Page 10 and Fig 1; 34pp; German.

RL JLT 24  
 AV 6253  
 ID AAR6253 standard; protein; 393 AA.  
 XA AAR6253;  
 XA AAR6253;  
 D 25-MAR-2003 (revised)  
 D 17-AUG-1995 (first entry)  
 D Bifunctional urokinase variant M20.  
 X fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
 K urokinase; variant; mutein.  
 X Synthetic.  
 X  
 F Key Location/Qualifiers  
 F Region 1..365  
 F /label= M4  
 F /note= "unglycosylated prourokinase(Ser47-Leu411)"  
 F Disulfide-bond 4..85  
 F Disulfide-bond 25..67  
 F Disulfide-bond 56..80  
 F Disulfide-bond 102..233  
 F Disulfide-bond 143..159  
 F Disulfide-bond 151..222  
 F Disulfide-bond 247..316  
 F Disulfide-bond 279..295  
 F Disulfide-bond 306..334  
 F Disulfide-bond 366..372  
 F Region /label= X1  
 F Region 373..393  
 F /label= Y1  
 DE4323754-C1.  
 01-DEC-1994.  
 15-JUL-1993; 93DE-04323754.  
 15-JUL-1993; 93DE-04323754.  
 (CHEF ) GRUENENTHAL GMBH.  
 Steffens GJ, Wnendt S, Schneider J, Heinzl-Wieland R;  
 Saunders DJ;  
 WPI; 1995-015191/03.  
 New bifunctional urokinase derivs and related plasmids - with improved  
 fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
 cerebral infarct, pulmonary embolism, etc.  
 Example 1; Page 10 and Fig 1; 34pp; German.  
 Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
 are claimed (see features table). Sequences AAR6244-R66266 are specific  
 examples of such derivs. which have both improved fibrinolytic and  
 thrombin-inhibiting activities, compared to known plasminogen activators  
 or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
 e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
 and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
 to correct PN field.)  
 Sequence 393 AA;  
 Query Match 100.0%; Score 510; DB 2; Length 393;  
 Best Local Similarity 100.0%; Pred. No. 8.8e-41;  
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 KTCYEGNGHYRGKASTDTWGRPCLPWNSATVLTQTYHAHRSDALQLGIGKINYCRNPDN 60  
 2 KTCYEGNGHYRGKASTDTWGRPCLPWNSATVLTQTYHAHRSDALQLGIGKINYCRNPDN 61

QY 61 RRRPWCYVQVGLKPLVQECWVHDCADGK 88  
 DB 62 RRRPWCYVQVGLKPLVQECWVHDCADGK 89  
 RESULT 25  
 AAR66249  
 ID AAR66249 standard; protein; 393 AA.  
 XX AAR66249;  
 AC AAR66249;  
 DT 25-MAR-2003 (revised)  
 DT 17-AUG-1995 (first entry)  
 XX Bifunctional urokinase variant M16.  
 XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
 KW urokinase; variant; mutein.  
 XX Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 XX Region 1..365  
 XX /label= M4  
 XX /note= "unglycosylated prourokinase(Ser47-Leu411)"  
 XX Disulfide-bond 4..85  
 XX Disulfide-bond 25..67  
 XX Disulfide-bond 56..80  
 XX Disulfide-bond 102..233  
 XX Disulfide-bond 143..159  
 XX Disulfide-bond 151..222  
 XX Disulfide-bond 247..316  
 XX Disulfide-bond 279..295  
 XX Disulfide-bond 306..334  
 XX Disulfide-bond 366..372  
 XX Region /label= X1  
 XX Region 373..393  
 XX /label= Y1  
 DE4323754-C1.  
 01-DEC-1994.  
 15-JUL-1993; 93DE-04323754.  
 15-JUL-1993; 93DE-04323754.  
 (CHEF ) GRUENENTHAL GMBH.  
 Steffens GJ, Wnendt S, Schneider J, Heinzl-Wieland R;  
 Saunders DJ;  
 WPI; 1995-015191/03.  
 New bifunctional urokinase derivs and related plasmids - with improved  
 fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
 cerebral infarct, pulmonary embolism, etc.  
 Example 1; Page 10 and Fig 1; 34pp; German.  
 Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
 are claimed (see features table). Sequences AAR6244-R66266 are specific  
 examples of such derivs. which have both improved fibrinolytic and  
 thrombin-inhibiting activities, compared to known plasminogen activators  
 or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
 e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
 and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
 to correct PN field.)  
 Sequence 393 AA;  
 Query Match 100.0%; Score 510; DB 2; Length 393;  
 1 KTCYEGNGHYRGKASTDTWGRPCLPWNSATVLTQTYHAHRSDALQLGIGKINYCRNPDN 60  
 2 KTCYEGNGHYRGKASTDTWGRPCLPWNSATVLTQTYHAHRSDALQLGIGKINYCRNPDN 61

KW urokinase; variant; mutein.  
 OX Synthetic.  
 OX  
 FT Key Location/Qualifiers  
 FT Region 1..365  
 FT /label= M4  
 FT /note= "unglycosylated prourokinase (Ser47-Leu411)"  
 FT Disulfide-bond 4..85  
 FT Disulfide-bond 25..67  
 FT Disulfide-bond 56..80  
 FT Disulfide-bond 102..233  
 FT Disulfide-bond 143..159  
 FT Disulfide-bond 151..222  
 FT Disulfide-bond 247..316  
 FT Disulfide-bond 279..295  
 FT Disulfide-bond 306..334  
 FT Disulfide-bond 366..371  
 FT Region /label= X1  
 FT Region 372..392  
 FT /label= Y1  
 FT DE4323754-Cl.  
 FT 01-DEC-1994.  
 FT 15-JUL-1993; 93DE-04323754.  
 FT 15-JUL-1993; 93DE-04323754.  
 FT (CHEF ) GRUENTHAL GMBH.  
 FT Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
 FT Saunders DJ;  
 FT WPI; 1995-015191/03.  
 FT New bifunctional urokinase derivs and related plasmids - with improved  
 FT fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
 FT cerebral infarct, pulmonary embolism, etc.  
 FT Example 1; Page 11 and Fig 1; 34pp; German.  
 FT Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
 FT are claimed (see features table). Sequences AAR6244-R6266 are specific  
 FT examples of such derivs. which have both improved fibrinolytic and  
 FT thrombin-inhibiting activities, compared to known plasminogen activators  
 FT or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
 FT e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
 FT and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
 FT to correct PN field.)  
 FT Sequence 392 AA;  
 FT Query Match 100.0%; Score 510; DB 2; Length 392;  
 FT Best Local Similarity 100.0%; Pred. No. 8.8e-41;  
 FT Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRPN 60  
 DY 2 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRPN 61  
 OY 61 RRRPWCYVQVGLKPLVQECWHDCA DGK 88  
 DY 62 RRRPWCYVQVGLKPLVQECWHDCA DGK 89  
 RF LT 23  
 AJ .6251  
 II AAR66251 standard; protein; 393 AA.  
 XZ  
 AC AAR66251;  
 XZ

DT 25-MAR-2003 (revised)  
 DT 17-AUG-1995 (first entry)  
 XX Bifunctional urokinase variant M18.  
 DE  
 XX fibrinolytic; thrombin inhibition; thrombolytic; anti-thrombotic;  
 KW urokinase; variant; mutein.  
 KW  
 XX Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 XX Region 1..365  
 XX /label= M4  
 XX /note= "unglycosylated prourokinase (Ser47-Leu411)"  
 XX Disulfide-bond 4..85  
 XX Disulfide-bond 25..67  
 XX Disulfide-bond 56..80  
 XX Disulfide-bond 102..233  
 XX Disulfide-bond 143..159  
 XX Disulfide-bond 151..222  
 XX Disulfide-bond 247..316  
 XX Disulfide-bond 279..295  
 XX Disulfide-bond 306..334  
 XX Disulfide-bond 366..372  
 XX Region /label= X1  
 XX Region 373..393  
 XX /label= Y1  
 XX DE4323754-Cl.  
 XX 01-DEC-1994.  
 XX 15-JUL-1993; 93DE-04323754.  
 XX 15-JUL-1993; 93DE-04323754.  
 XX (CHEF ) GRUENTHAL GMBH.  
 XX Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
 XX Saunders DJ;  
 XX WPI; 1995-015191/03.  
 XX New bifunctional urokinase derivs and related plasmids - with improved  
 XX fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
 XX cerebral infarct, pulmonary embolism, etc.  
 XX Example 1; Page 10 and Fig 1; 34pp; German.  
 XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
 XX are claimed (see features table). Sequences AAR6244-R6266 are specific  
 XX examples of such derivs. which have both improved fibrinolytic and  
 XX thrombin-inhibiting activities, compared to known plasminogen activators  
 XX or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
 XX e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
 XX and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
 XX to correct PN field.)  
 XX Sequence 393 AA;  
 SQ Query Match 100.0%; Score 510; DB 2; Length 393;  
 SQ Best Local Similarity 100.0%; Pred. No. 8.8e-41;  
 SQ Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRPN 60  
 DY 2 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRPN 61  
 OY 61 RRRPWCYVQVGLKPLVQECWHDCA DGK 88  
 DY 62 RRRPWCYVQVGLKPLVQECWHDCA DGK 89

F Disulfide-bond 143. .159  
F Disulfide-bond 151. .222  
F Disulfide-bond 247. .316  
F Disulfide-bond 279. .295  
F Disulfide-bond 306. .334  
F Region 366. .371  
F /label= X1  
F Region 372. .392  
F /label= Y1  
X DE4323754-C1.  
X 01-DEC-1994.  
X 15-JUL-1993; 93DE-04323754.  
X 15-JUL-1993; 93DE-04323754.  
X (CHEF ) GRUENTHAL GMBH.  
X Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
X Saunders DU;  
X WPI; 1995-015191/03.  
X New bifunctional urokinase derivs and related plasmids - with improved  
X fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
X cerebral infarct, pulmonary embolism, etc.  
X Example 1; Page 10 and Fig 1; 34pp; German.  
X Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
X are claimed (see features table). Sequences AAR66244-R66266 are specific  
X examples of such derivs. which have both improved fibrinolytic and  
X thrombin-inhibiting activities, compared to known plasminogen activators  
X or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
X e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
X and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
X to correct PN field.)  
X Sequence-392 AA;  
X Query Match 100.0%; Score 510; DB 2; Length 392;  
X Best Local Similarity 100.0%; Pred. No. 8.8e-41;  
X Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Q 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRS DALQLGLGKHNYCRNPDN 60  
D 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRS DALQLGLGKHNYCRNPDN 61  
Q 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
D 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89  
R JLT 21  
A 56256  
I AAR66256 standard; protein; 392 AA.  
X AAR66256;  
X 25-MAR-2003 (revised)  
D 17-AUG-1995 (first entry)  
X Bifunctional urokinase variant M23.  
X fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
X urokinase; variant; mutein.  
X Synthetic.  
X Key Location/Qualifiers  
F Region 1. .365

FT /label= M4  
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"  
FT 4. .85  
FT Disulfide-bond 25. .67  
FT Disulfide-bond 56. .80  
FT Disulfide-bond 102. .233  
FT Disulfide-bond 143. .159  
FT Disulfide-bond 151. .222  
FT Disulfide-bond 247. .316  
FT Disulfide-bond 279. .295  
FT Disulfide-bond 306. .334  
FT Region 366. .371  
FT /label= X1  
FT Region 372. .392  
FT /label= Y1  
XX DE4323754-C1.  
XX 01-DEC-1994.  
XX 15-JUL-1993; 93DE-04323754.  
XX 15-JUL-1993; 93DE-04323754.  
XX (CHEF ) GRUENTHAL GMBH.  
XX Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
XX Saunders DU;  
XX WPI; 1995-015191/03.  
XX New bifunctional urokinase derivs and related plasmids - with improved  
XX fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
XX cerebral infarct, pulmonary embolism, etc.  
XX Example 1; Page 10 and Fig 1; 34pp; German.  
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
XX are claimed (see features table). Sequences AAR66244-R66266 are specific  
XX examples of such derivs. which have both improved fibrinolytic and  
XX thrombin-inhibiting activities, compared to known plasminogen activators  
XX or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
XX e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
XX and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
XX to correct PN field.)  
XX SQ Sequence 392 AA;  
XX Query Match 100.0%; Score 510; DB 2; Length 392;  
XX Best Local Similarity 100.0%; Pred. No. 8.8e-41;  
XX Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRS DALQLGLGKHNYCRNPDN 60  
DB 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRS DALQLGLGKHNYCRNPDN 61  
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89  
RESULT 22  
AAR66257  
ID AAR66257 standard; protein; 392 AA.  
XX AAR66257;  
XX 25-MAR-2003 (revised)  
DT 17-AUG-1995 (first entry)  
XX Bifunctional urokinase variant M24.  
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;

P/ 01-DEC-1994.  
X/ 15-JUL-1993; 93DE-04323754.  
P/ 15-JUL-1993; 93DE-04323754.  
X/ (CHEF ) GRUENENTHAL GMBH.  
P/ Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
P/ Saunders DJ;  
D/ WPI; 1995-015191/03.  
X/ New bifunctional urokinase derive and related plasmids - with improved  
P/ fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
P/ cerebral infarct, pulmonary embolism, etc.  
X/ Example 1; Page 11 and Fig 1; 34pp; German.  
X/ Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
C/ are claimed (see features table). Sequences AAR66244-R66266 are specific  
C/ examples of such derive, which have both improved fibrinolytic and  
C/ thrombin-inhibiting activities, compared to known plasminogen activators  
C/ or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
C/ e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
C/ and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
C/ to correct PN field.)  
X/ Sequence 392 AA;  
S/ ery Match 100.0%; Score 510; DB 2; Length 392;  
st Local Similarity 100.0%; Pred. No. 8.8e-41;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Q/ 1 KTCYEGNGHFRGKASTDTMGRCPLPWSATVLOQTYHAHRSALQGLGKHNYCRNPDN 60  
D/ 2 KTCYEGNGHFRGKASTDTMGRCPLPWSATVLOQTYHAHRSALQGLGKHNYCRNPDN 61  
Q/ 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
D/ 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89  
R/ LT 19  
A/ 6263  
I/ AAR66263 standard; protein; 392 AA.  
X/ AAR66263;  
A/ 25-MAR-2003 (revised)  
D/ 17-AUG-1995 (first entry)  
X/ Bifunctional urokinase variant M30.  
X/ fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
X/ urokinase; variant; mutein.  
O/ Synthetic.  
F/ Key Location/Qualifiers  
F/ Region 1..365  
F/ /label= M4  
F/ /note= "unglycosylated prourokinase (Ser47-Leu411)"  
F/ Disulfide-bond 4..85  
F/ Disulfide-bond 25..67  
F/ Disulfide-bond 56..80  
F/ Disulfide-bond 102..233  
F/ Disulfide-bond 143..159  
F/ Disulfide-bond 151..222  
F/ Disulfide-bond 247..316  
F/ Disulfide-bond 279..295  
F/ Disulfide-bond 306..334  
F/ Disulfide-bond 366..371  
F/ Region

FT Region /label= X1  
FT 372..392  
XX /label= Y1  
PN DE4323754-C1.  
XX  
PD 01-DEC-1994.  
XX  
PF 15-JUL-1993; 93DE-04323754.  
XX  
PR 15-JUL-1993; 93DE-04323754.  
XX  
PA (CHEF ) GRUENENTHAL GMBH.  
XX  
PI Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
PI Saunders DJ;  
XX  
DR WPI; 1995-015191/03.  
XX  
XX New bifunctional urokinase derivs and related plasmids - with improved  
PT fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
PT cerebral infarct, pulmonary embolism, etc.  
XX  
PS Example 1; Page 11 and Fig 1; 34pp; German.  
XX  
CC Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
CC are claimed (see features table). Sequences AAR66244-R66266 are specific  
CC examples of such derivs, which have both improved fibrinolytic and  
CC thrombin-inhibiting activities, compared to known plasminogen activators  
CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
CC to correct PN field.)  
XX  
SQ Sequence 392 AA;  
Query Match 100.0%; Score 510; DB 2; Length 392;  
Best Local Similarity 100.0%; Pred. No. 8.8e-41;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHFRGKASTDTMGRCPLPWSATVLOQTYHAHRSALQGLGKHNYCRNPDN 60  
Db 2 KTCYEGNGHFRGKASTDTMGRCPLPWSATVLOQTYHAHRSALQGLGKHNYCRNPDN 61  
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
Db 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89  
RESULT 20  
AAR66254  
ID AAR66254 standard; protein; 392 AA.  
XX  
AC AAR66254;  
XX  
XX 25-MAR-2003 (revised)  
DT 17-AUG-1995 (first entry)  
XX  
XX Bifunctional urokinase variant M21.  
DE fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
XX urokinase; variant; mutein.  
KW Synthetic.  
XX  
XX Key Location/Qualifiers  
FH Region 1..365  
FT /label= M4  
FT /note= "unglycosylated prourokinase (Ser47-Leu411)"  
FT Disulfide-bond 4..85  
FT Disulfide-bond 25..67  
FT Disulfide-bond 56..80  
FT Disulfide-bond 102..233

XX New bifunctional urokinase derivs and related plasmids - with improved  
XX fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
XX cerebral infarct, pulmonary embolism, etc.  
XX Example 1; Page 11 and Fig 1; 34pp; German.  
XX  
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
XX are claimed (see features table). Sequences AAR6244-R6266 are specific  
XX examples of such derivs which have both improved fibrinolytic and  
XX thrombin-inhibiting activities, compared to known plasminogen activators  
XX or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
XX e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
XX and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
XX to correct PN field.)  
XX Sequence 392 AA;  
XX  
XX Query Match 100.0%; Score 510; DB 2; Length 392;  
XX Best Local Similarity 100.0%; Pred. No. 8.8e-41;  
XX Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 1. KTCYEGNGHFYRGKASTDTMGRCPLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNPDN 60  
XX 2. KTCYEGNGHFYRGKASTDTMGRCPLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNPDN 61  
XX  
XX 61. RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
XX 62. RRRPWCYVQVGLKPLVQECMVHDCADGK 89  
XX  
XX LT 17  
XX A. 6261  
XX AAR66261 standard; protein; 392 AA.  
XX AAR66261;  
XX  
XX 25-MAR-2003 (revised)  
XX 17-AUG-1995 (first entry)  
XX  
XX Bifunctional urokinase variant M28.  
XX  
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
XX urokinase; variant; mutein.  
XX  
XX Synthetic.  
XX  
XX Key Location/Qualifiers  
XX Region 1. 365  
XX /label= M4  
XX /note= "unglycosylated prourokinase (Ser47-Leu411)"  
XX  
XX Disulfide-bond 25. 67  
XX Disulfide-bond 56. 80  
XX Disulfide-bond 102. 233  
XX Disulfide-bond 143. 159  
XX Disulfide-bond 151. 222  
XX Disulfide-bond 247. 316  
XX Disulfide-bond 279. 295  
XX Disulfide-bond 306. 334  
XX Region 366. 371  
XX /label= X1  
XX Region 372. 392  
XX /label= Y1  
XX  
XX DE4323754-C1.  
XX  
XX 01-DEC-1994.  
XX  
XX 15-JUL-1993; 93DE-04323754.  
XX  
XX 15-JUL-1993; 93DE-04323754.  
XX

PA (CHEF ) GRUENTHAL GMBH.  
XX Steffens GJ, Wnendt S, Schneider J, Heinzl-Wieland R;  
PI Saunders DJ;  
XX  
XX WPI; 1995-015191/03.  
XX  
XX New bifunctional urokinase derivs and related plasmids - with improved  
XX fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
XX cerebral infarct, pulmonary embolism, etc.  
XX Example 1; Page 11 and Fig 1; 34pp; German.  
XX  
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
XX are claimed (see features table). Sequences AAR6244-R6266 are specific  
XX examples of such derivs which have both improved fibrinolytic and  
XX thrombin-inhibiting activities, compared to known plasminogen activators  
XX or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
XX e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
XX and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
XX to correct PN field.)  
XX Sequence 392 AA;  
XX  
XX Query Match 100.0%; Score 510; DB 2; Length 392;  
XX Best Local Similarity 100.0%; Pred. No. 8.8e-41;  
XX Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 1. KTCYEGNGHFYRGKASTDTMGRCPLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNPDN 60  
XX 2. KTCYEGNGHFYRGKASTDTMGRCPLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNPDN 61  
XX  
XX 61. RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
XX 62. RRRPWCYVQVGLKPLVQECMVHDCADGK 89  
XX  
XX  
XX AAR66258 standard; protein; 392 AA.  
XX  
XX AAR66258;  
XX  
XX 25-MAR-2003 (revised)  
XX 17-AUG-1995 (first entry)  
XX  
XX Bifunctional urokinase variant M25.  
XX  
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
XX urokinase; variant; mutein.  
XX  
XX Synthetic.  
XX  
XX Key Location/Qualifiers  
XX Region 1. 365  
XX /label= M4  
XX /note= "unglycosylated prourokinase (Ser47-Leu411)"  
XX  
XX Disulfide-bond 4. 85  
XX Disulfide-bond 25. 67  
XX Disulfide-bond 56. 80  
XX Disulfide-bond 102. 233  
XX Disulfide-bond 143. 159  
XX Disulfide-bond 151. 222  
XX Disulfide-bond 247. 316  
XX Disulfide-bond 279. 295  
XX Disulfide-bond 306. 334  
XX Region 366. 371  
XX /label= X1  
XX Region 372. 392  
XX /label= Y1  
XX  
XX DE4323754-C1.  
XX

C e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
C and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
C to correct PN field.)  
X Sequence 392 AA;  
S Query Match 100.0%; Score 510; DB 2; Length 392;  
st Local Similarity 100.0%; Pred. No. 8.8e-41;  
tches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Q 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNCRNPDN 60  
D 2 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNCRNPDN 61  
Q 61 RRRPWCYVQVGLKPLVQECMWHDCADGK 88  
D 62 RRRPWCYVQVGLKPLVQECMWHDCADGK 89  
RF LT 15  
A .6255  
I AAR66255 standard; protein; 392 AA.  
X AAR66255;  
D 25-MAR-2003 (revised)  
D 17-AUG-1995 (first entry)  
X Bifunctional urokinase variant M22.  
X fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
K urokinase; variant; mutein.  
X Synthetic.  
X Key Location/Qualifiers  
F Region 1..365  
F /label= M4  
F /note= "unglycosylated prourokinase(Ser47-Leu411)"  
F Disulfide-bond 4..85  
F Disulfide-bond 25..67  
F Disulfide-bond 56..80  
F Disulfide-bond 102..233  
F Disulfide-bond 143..159  
F Disulfide-bond 151..222  
F Disulfide-bond 247..316  
F Disulfide-bond 279..295  
F Disulfide-bond 306..334  
F Region 366..371  
F /label= X1  
F Region 372..392  
F /label= Y1  
D DE4323754-C1.  
X 01-DEC-1994.  
X 15-JUL-1993; 93DE-04323754.  
X 15-JUL-1993; 93DE-04323754.  
X (CHEF ) GRUENTHAL GMBH.  
X Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
P Saunders DJ;  
X WPI; 1995-015191/03.  
D New bifunctional urokinase derivs and related plasmids - with improved  
P fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
P cerebral infarct, pulmonary embolism, etc.  
X Example 1; Page 10 and Fig 1; 34pp; German.

XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
CC are claimed (see features table). Sequences AAR66244-R66266 are specific  
CC examples of such derivs. which have both improved fibrinolytic and  
CC thrombin-inhibiting activities, compared to known plasminogen activators  
CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
CC to correct PN field.)  
X Sequence 392 AA;  
S Query Match 100.0%; Score 510; DB 2; Length 392;  
Best Local Similarity 100.0%; Pred. No. 8.8e-41;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Q 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNCRNPDN 60  
D 2 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNCRNPDN 61  
Q 61 RRRPWCYVQVGLKPLVQECMWHDCADGK 88  
D 62 RRRPWCYVQVGLKPLVQECMWHDCADGK 89  
RF LT 16  
A AAR66259 standard; protein; 392 AA.  
X AAR66259;  
D 25-MAR-2003 (revised)  
D 17-AUG-1995 (first entry)  
X Bifunctional urokinase variant M26.  
X fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
K urokinase; variant; mutein.  
X Synthetic.  
X Key Location/Qualifiers  
F Region 1..365  
F /label= M4  
F /note= "unglycosylated prourokinase (Ser47-Leu411)"  
F Disulfide-bond 4..85  
F Disulfide-bond 25..67  
F Disulfide-bond 56..80  
F Disulfide-bond 102..233  
F Disulfide-bond 143..159  
F Disulfide-bond 151..222  
F Disulfide-bond 247..316  
F Disulfide-bond 279..295  
F Disulfide-bond 306..334  
F Region 366..371  
F /label= X1  
F Region 372..392  
F /label= Y1  
D DE4323754-C1.  
X 01-DEC-1994.  
X 15-JUL-1993; 93DE-04323754.  
X 15-JUL-1993; 93DE-04323754.  
X (CHEF ) GRUENTHAL GMBH.  
X Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
P Saunders DJ;  
X WPI; 1995-015191/03.



DI 2 KTCYEGNGHFYRGKASTDTMGRCPLPWNSATVLTQTYHAHRSALQLGLGKHNYCRNPND 61  
Q 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
DI 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89  
R 13  
A 6260  
X AAR66260 standard; protein; 392 AA.  
X AAR66260;  
X  
D 25-MAR-2003 (revised)  
D 17-AUG-1995 (first entry)  
X Bifunctional urokinase variant M27.  
X fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
X urokinase; variant; mutein.  
X Synthetic.  
F Key Location/Qualifiers  
F Region 1..365  
F /label= M4  
F /note= "unglycosylated prourokinase(Ser47-Leu411)"  
F Disulfide-bond 4..85  
F Disulfide-bond 25..67  
F Disulfide-bond 56..80  
F Disulfide-bond 102..233  
F Disulfide-bond 143..159  
F Disulfide-bond 151..222  
F Disulfide-bond 247..316  
F Disulfide-bond 279..295  
F Disulfide-bond 306..334  
F Disulfide-bond 366..371  
F Region /label= X1  
F Region 372..392  
F /label= Y1  
X DE4323754-Cl  
X 01-DEC-1994.  
X 15-JUL-1993; 93DE-04323754.  
X 15-JUL-1993; 93DE-04323754.  
X (CHEF ) GRUENENTHAL GMBH.  
X Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
X Saunders DJ;  
X WPI; 1995-015191/03.  
X New bifunctional urokinase derivs and related plasmids - with improved  
X fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
X cerebral infarct, pulmonary embolism, etc.  
X Example 1; Page 11 and Fig 1; 34pp; German.  
X Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
X are claimed (see features table). Sequences AAR66264-R66266 are specific  
X examples of such derivs, which have both improved fibrinolytic and  
X thrombin-inhibiting activities, compared to known plasminogen activators  
X or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
X e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
X and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
X to correct PN field.)  
X Sequence 392 AA;

Query Match 100.0%; Score 510; DB 2; Length 392;  
Best Local Similarity 100.0%; Pred. NO. 8.8e-41;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWNSATVLTQTYHAHRSALQLGLGKHNYCRNPND 60  
DB 2 KTCYEGNGHFYRGKASTDTMGRCPLPWNSATVLTQTYHAHRSALQLGLGKHNYCRNPND 61  
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89  
RESULT 14  
AAR66264  
ID AAR66264 standard; protein; 392 AA.  
XX  
AC AAR66264;  
XX  
DT 25-MAR-2003 (revised)  
DT 17-AUG-1995 (first entry)  
XX  
DE Bifunctional urokinase variant M31.  
XX  
KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
KW urokinase; variant; mutein.  
XX  
OS Synthetic.  
PH Key Location/Qualifiers  
FT Region 1..365  
FT /label= M4  
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"  
FT Disulfide-bond 4..85  
FT Disulfide-bond 25..67  
FT Disulfide-bond 56..80  
FT Disulfide-bond 102..233  
FT Disulfide-bond 143..159  
FT Disulfide-bond 151..222  
FT Disulfide-bond 247..316  
FT Disulfide-bond 279..295  
FT Disulfide-bond 306..334  
FT Region /label= X1  
FT Region 372..392  
FT /label= Y1  
XX DE4323754-Cl.  
XX  
XX 01-DEC-1994.  
XX  
XX 15-JUL-1993; 93DE-04323754.  
XX PR 15-JUL-1993; 93DE-04323754.  
XX (CHEF ) GRUENENTHAL GMBH.  
XX Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
XX Saunders DJ;  
XX WPI; 1995-015191/03.  
XX  
XX New bifunctional urokinase derivs and related plasmids - with improved  
XX fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
XX cerebral infarct, pulmonary embolism, etc.  
XX Example 1; Page 11 and Fig 1; 34pp; German.  
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
XX are claimed (see features table). Sequences AAR66264-R66266 are specific  
XX examples of such derivs, which have both improved fibrinolytic and  
XX thrombin-inhibiting activities, compared to known plasminogen activators  
XX or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
XX e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
XX and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
XX to correct PN field.)  
XX Sequence 392 AA;

```
XJ 25-MAR-2003 (revised)
D1 17-AUG-1995 (first entry)
XJ Bifunctional urokinase variant M12.
XJ fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
XJ urokinase; variant; mutein.
XJ Synthetic.
XJ
XJ Key Location/Qualifiers
XJ Region 1..365
XJ /label= M4
XJ /note= "unglycosylated prourokinase (Ser47-Leu411)"
XJ Disulfide-bond 25..67
XJ Disulfide-bond 56..80
XJ Disulfide-bond 102..233
XJ Disulfide-bond 143..159
XJ Disulfide-bond 151..222
XJ Disulfide-bond 247..316
XJ Disulfide-bond 279..295
XJ Disulfide-bond 306..334
XJ Disulfide-bond 366..371
XJ Region /label= X1
XJ Region /label= Y1
XJ DE4323754-C1.
XJ 01-DEC-1994.
XJ 15-JUL-1993; 93DE-04323754.
XJ 15-JUL-1993; 93DE-04323754.
XJ (CHEF ) GRUENENTHAL GMBH.
XJ Steffens GJ, Wrendt S, Schneider J, Heinzel-Wieland R;
XJ Saunders DJ;
XJ WPI; 1995-015191/03.
XJ New bifunctional urokinase derivs and related plasmids - with improved
XJ fibrinolytic and thrombin inhibiting activities, for treating cardiac and
XJ cerebral infarct, pulmonary embolism, etc.
XJ Example 1; Page 10 and Fig 1; 34pp; German.
XJ Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1
XJ are claimed (see features table). Sequences AAR66244-R66266 are specific
XJ examples of such derivs. which have both improved fibrinolytic and
XJ thrombin-inhibiting activities, compared to known plasminogen activators
XJ or thrombin inhibitors. The proteins are useful as thrombolytic agents,
XJ e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac
XJ and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003
XJ to correct PN field.)
XJ Sequence 390 AA;
XJ
XJ Query Match 100.0%; Score 510; DB 2; Length 390;
XJ StLocal Similarity 100.0%; Pred. No. 8.8e-41;
XJ Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XJ
XJ 1 KTCYEGNGHFRGKASTDTMGRCPLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
XJ 2 KTCYEGNGHFRGKASTDTMGRCPLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61
XJ 61 RRRPWCYVQVGLKPLVQECVHDCADGK 88
XJ 62 RRRPWCYVQVGLKPLVQECVHDCADGK 89
XJ
```

```
RESULT 12
AAR66247
ID AAR66247 standard; protein; 390 AA.
XX
XX AAR66247;
XX
XX 25-MAR-2003 (revised)
DT 17-AUG-1995 (first entry)
XX
XX Bifunctional urokinase variant M14.
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
XX urokinase; variant; mutein.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Region 1..365
XX /label= M4
XX /note= "unglycosylated prourokinase (Ser47-Leu411)"
XX Disulfide-bond 25..67
XX Disulfide-bond 56..80
XX Disulfide-bond 102..233
XX Disulfide-bond 143..159
XX Disulfide-bond 151..222
XX Disulfide-bond 247..316
XX Disulfide-bond 279..295
XX Disulfide-bond 306..334
XX Disulfide-bond 366..371
XX Region /label= X1
XX Region /label= Y1
XX DE4323754-C1.
XX 01-DEC-1994.
XX 15-JUL-1993; 93DE-04323754.
XX 15-JUL-1993; 93DE-04323754.
XX (CHEF ) GRUENENTHAL GMBH.
XX Steffens GJ, Wrendt S, Schneider J, Heinzel-Wieland R;
XX Saunders DJ;
XX WPI; 1995-015191/03.
XX New bifunctional urokinase derivs and related plasmids - with improved
XX fibrinolytic and thrombin inhibiting activities, for treating cardiac and
XX cerebral infarct, pulmonary embolism, etc.
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1
XX are claimed (see features table). Sequences AAR66244-R66266 are specific
XX examples of such derivs. which have both improved fibrinolytic and
XX thrombin-inhibiting activities, compared to known plasminogen activators
XX or thrombin inhibitors. The proteins are useful as thrombolytic agents,
XX e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac
XX and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003
XX to correct PN field.)
XX Sequence 390 AA;
XX
XX Query Match 100.0%; Score 510; DB 2; Length 390;
XX Best Local Similarity 100.0%; Pred. No. 8.8e-41;
XX Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 KTCYEGNGHFRGKASTDTMGRCPLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
XX
```

Q: Synthetic.  
X: Key  
F: Region  
F: Location/Qualifiers  
F: 1.365  
F: /label= M4  
F: /note= "unglycosylated prourokinase(Ser47-Leu411)"  
F: 4.85  
F: Disulfide-bond 25.67  
F: Disulfide-bond 56.80  
F: Disulfide-bond 102.233  
F: Disulfide-bond 143.159  
F: Disulfide-bond 151.222  
F: Disulfide-bond 247.316  
F: Disulfide-bond 279.295  
F: Disulfide-bond 306.334  
F: Disulfide-bond 365.366  
F: Region  
F: /label= X1  
F: /note= "peptide bond"  
F: 366.386  
F: /label= Y1  
X: DB4323754-C1  
X: 01-DEC-1994  
X: 15-JUL-1993; 93DE-04323754.  
X: 15-JUL-1993; 93DE-04323754.  
X: (CHEF ) GRUENTHAL GMBH.  
X: Steffens GJ, Whendt S, Schneider J, Heinzl-Wieland R;  
X: Saunders DU;  
X: WPI; 1995-015191/03.  
X: New bifunctional urokinase derivs and related plasmids - with improved  
X: fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
X: cerebral infarct, pulmonary embolism, etc.  
X: Example 1; Page 11 and Fig 1; 34pp; German.  
X: Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
X: are claimed (see features table). Sequences AAR66244-R66266 are specific  
X: examples of such derivs which have both improved fibrinolytic and  
X: thrombin-inhibiting activities, compared to known plasminogen activators  
X: or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
X: e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
X: and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
X: to correct PN field.)  
X: Sequence 386 AA;  
X: Query Match 100.0%; Score 510; DB 2; Length 386;  
X: Best Local Similarity 100.0%; Pred. No. 8.7e-41;  
X: Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Q: 1 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNP DN 60  
D: 2 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNP DN 61  
Q: 61 RRRPWCYVQVGLKPLVQECWVHDCADGK 88  
D: 62 RRRPWCYVQVGLKPLVQECWVHDCADGK 89  
R: 1LT 10  
A: 3636  
I: AAW13636 standard; protein; 389 AA.  
X: AAW13636;  
X: 04-JUN-1997 (first entry)

XX Human prourokinase variant lacking EGF domain loops 1 and 2.  
DE  
XX  
XX Human; prourokinase; hPUK; variant; half-life; increase; EGF;  
KW epidermal growth factor domain; deletion; thrombolysis; fibrinolysis.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
XX Location/Qualifiers  
XX Key  
XX Region 1..10  
XX /note= "residues 1-10 of native hPUK"  
FT Region 11..389  
FT /note= "residues 33-411 of native hPUK"  
FT Misc-difference 155  
FT /note= "corresponds to TAC codon"  
XX  
XX EP398361-A.  
XX 22-NOV-1990.  
XX  
XX 18-MAY-1990; 90EP-00109472.  
XX  
XX 18-MAY-1989; 89JP-00126433.  
XX 22-FEB-1990; 90JP-00042020.  
XX  
XX (GRC ) GREEN CROSS CORP.  
XX Tanabe T, Amatsuji Y, Kasai S, Hirose M, Morita M, Kawabe H;  
XX Airmura H;  
XX WPI; 1990-350146/47.  
XX N-PSDB; AA161673.  
XX  
XX Human pro-urokinase variants - deficient in loop regions of epidermal  
XX growth factor, showing long blood half-life, as fibrinolytic agent.  
XX Claim 6; Page; 22pp; English.  
XX  
XX New variants of human prourokinase (hPUK) comprise a hPUK deficient in  
XX (i) at least part of the first loop region of the epidermal growth factor  
XX (EGF) domain; (ii) at least part of the first loop and at least part of  
XX the second loop; or (iii) at least part of the third loop. The hPUK  
XX variants show an increased blood half-life comparable to that of the  
XX whole EGF domain-deficient hPUK variant and urokinase while retaining the  
XX same properties as those of hPUK. They have potent thrombolytic activity  
XX and very little tendency to cause spontaneous bleeding. The present  
XX sequence represents a specific variant of hPUK which lacks loops 1 and 2  
XX of the EGF domain; the sequence does not appear in the specification and  
XX has been created using the wild-type hPUK sequence and the junction  
XX sequence after deletion, both of which are given (in Fig 1 and on page 8,  
XX respectively)  
XX  
XX Sequence 389 AA;  
XX  
XX Query Match 100.0%; Score 510; DB 2; Length 389;  
XX Best Local Similarity 100.0%; Pred. No. 8.7e-41;  
XX Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Q: 1 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNP DN 60  
D: 26 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNP DN 85  
Q: 61 RRRPWCYVQVGLKPLVQECWVHDCADGK 88  
D: 86 RRRPWCYVQVGLKPLVQECWVHDCADGK 113  
RESULT 11  
AAR66245  
ID AAR66245 standard; protein; 390 AA.  
XX  
XX AAR66245;  
AC

XJ Human, des-epidermal growth factor homologous plasminogen activator; uPA;  
 KQ liver membrane; reduced affinity; EGF homologous; thrombolytic;  
 KQ thrombolytic; increased half-life; urokinase.  
 XQ Homo sapiens; (engineered).  
 XJ Key Location/Qualifiers  
 FT Misc-difference 1  
 FT /note= "amino acids 1-46 of wild-type urokinase have been  
 FT deleted"  
 XQ US5376547-A.  
 XQ 27-DEC-1994.  
 XQ 29-JAN-1988; 88US-00150267.  
 XQ 30-JAN-1987; 87US-00008795.  
 XQ (AMGP ) AMERICAN HOME PROD CORP.  
 XQ Hung PP, Lee SL, Kalyan NK;  
 XQ WPI; 1995-043464/06.  
 XQ New modified plasminogen activator cpds. - having regions removed to  
 XQ reduce affinity for liver membranes and increase circulation half-life.  
 XQ Claim 1; Page ?; 26pp; English.  
 XQ Amino acid residues 1-46 contain the EGF region of human urokinase.  
 XQ Deletion of this region results in a plasminogen activator with reduced  
 XQ affinity for liver cell membranes; the mutant protein is not cleared from  
 XQ the circulation as rapidly as is wild-type tPA. The specification only  
 XQ gives the sequence around the deletion and not the full-length sequence  
 XQ of "delta 1-46 urokinase"; the sequence in AAR68854 has been obtained by  
 XQ amending a previously disclosed wild-type human urokinase sequence (from  
 XQ WO9501427) according to the description given in Example 3. (Updated on  
 XQ 25-MAR-2003 to correct PF field.) (Updated on 16-OCT-2003 to standardise  
 XQ OS field)  
 XQ Sequence 365 AA;  
 Query Match 100.0%; Score 510; DB 2; Length 365;  
 st Local Similarity 100.0%; Pred. No. 8.2e-41;  
 tches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KTCYEGNGHFGYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 60  
 DB 2 KTCYEGNGHFGYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 61  
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
 DL 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89  
 RE LT 8  
 AA 3635  
 II AAW13635 standard; protein; 378 AA.  
 XQ AAW13635;  
 XQ 04-JUN-1997 (first entry)  
 XQ Human prourokinase variant lacking entire EGF domain.  
 XQ Human; prourokinase; hPUK; variant; half-life; increase; EGF;  
 XQ epidermal growth factor domain; deletion; thrombolysis; fibrinolysis.  
 XQ Homo sapiens.  
 OS Synthetic.  
 XQ

PH Key Location/Qualifiers  
 FT Region 1..9  
 FT /note= "residues 1-9 of native hPUK"  
 FT Region 10..378  
 FT /note= "residues 43-411 of native hPUK"  
 FT Misc-difference 144  
 FT /note= "corresponds to TAC codon"  
 XX EP398361-A.  
 XX 22-NOV-1990.  
 XX 18-MAY-1990; 90EP-00109472.  
 XX 18-MAY-1989; 89JP-00126433.  
 XX 22-FEB-1990; 90JP-00042020.  
 XX (GREG ) GREEN CROSS CORP.  
 XX Tanabe T, Amatsuji Y, Kasai S, Hirose M, Morita M, Kawabe H;  
 XX Airmura H;  
 XX WPI; 1990-350146/47.  
 XX N-PSDB; AAT61672.  
 XX Human pro-urokinase variants - deficient in loop regions of epidermal  
 XX growth factor, showing long blood half-life, as fibrinolytic agent.  
 XX Claim 1; Page; 22pp; English.  
 XX New variants of human prourokinase (hPUK) comprise a hPUK deficient in  
 XX (i) at least part of the first loop region of the epidermal growth factor  
 XX (EGF) domain; (ii) at least part of the first loop and at least part of  
 XX the second loop; or (iii) at least part of the third loop. The hPUK  
 XX variants show an increased blood half-life comparable to that of the  
 XX whole EGF domain-deficient hPUK variant and urokinase while retaining the  
 XX same properties as those of hPUK. They have potent thrombolytic activity  
 XX and very little tendency to cause spontaneous bleeding. The present  
 XX sequence represents a specific variant of hPUK which lacks the entire EGF  
 XX domain; the sequence does not appear in the specification and has been  
 XX created using the wild-type hPUK sequence and the junction sequence after  
 XX deletion, both of which are given (in Fig 1 and in Fig 2(3),  
 XX respectively)  
 XX Sequence 378 AA;  
 Query Match 100.0%; Score 510; DB 2; Length 378;  
 Best Local Similarity 100.0%; Pred. No. 8.5e-41;  
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KTCYEGNGHFGYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 60  
 DB 15 KTCYEGNGHFGYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 74  
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
 DB 75 RRRPWCYVQVGLKPLVQECMVHDCADGK 102  
 RESULT 9  
 AAR66266  
 ID AAR66266 standard; protein; 386 AA.  
 XX AAR66266;  
 XX 25-MAR-2003 (revised)  
 DT 17-AUG-1995 (first entry)  
 XX Bifunctional urokinase variant M33.  
 XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
 XX urokinase; variant; mutein.  
 XX

X Ruben SM, Baraah SC, Birse CE, Rosen CA;  
 P WPI; 2001-235357/24.  
 D N-PSDB; AAB34897.  
 D  
 X Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
 P useful for preventing, diagnosing and/or treating colorectal cancers.  
 X  
 F Claim 11; Page 7707-7708; 9803pp; English.  
 X  
 X AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
 C cancer-associated nucleic acid molecules (N) and proteins (P), where the  
 C proteins are collectively known as colon cancer antigens. The colon  
 C cancer antigens have cytostatic activity and can be used in gene therapy  
 C and treatment of diseases associated with inappropriate P expression. For  
 C example, N and P may be used to treat disorders associated with decreased  
 C expression by rectifying mutations or deletions in a patient's genome  
 C that affect the activity of P by expressing inactive proteins or to  
 C supplement the patient's own production of P. Additionally, N may be used  
 C to produce the colon cancer-associated Pa, by inserting the nucleic acids  
 C into a host cell and culturing the cell to express the proteins. N and P  
 C can be used in the prevention, diagnosis and treatment of colorectal  
 C carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent  
 C sequences used in the exemplification of the present invention. N.B.  
 C Pages 666 to 682 and page 7053 of the sequence listing were missing at  
 C time of publication, meaning no sequences are present for SEQ ID NO:1027  
 C to 1052, 7921 and 7922  
 X  
 S Sequence 337.AA;

Query Match 100.0%; Score 510; DB 4; Length 337;  
 Best Local Similarity 100.0%; Pred. No. 7.6e-41;  
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KTCYEGNGHFYRGKASTDTMGPRCLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60  
 74 KTCYEGNGHFYRGKASTDTMGPRCLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 133  
 61 RRRPWCYVQVGLKPLVQECWVHDCADGK 88  
 134 RRRPWCYVQVGLKPLVQECWVHDCADGK 161

R ULT 6  
 A 11795  
 I ABP41795 standard; protein; 337 AA.  
 X  
 X ABP41795;  
 X  
 X 22-AUG-2002 (first entry)

X Human ovarian antigen HVVCB79, SEQ ID NO:2927.

X Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
 X ovarian cancer; breast cancer; tumour; reproductive system disorder;  
 X infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
 X PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;  
 X inflammatory condition; immune disorder; blood disorder;  
 X cardiovascular disorder; respiratory disorder; neurological disorder;  
 X gastrointestinal disorder; urinary system disorder; drug screening;  
 X gene therapy; chromosome mapping; forensic analysis;  
 X antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
 X antiinflammatory; gynaecological; reproductive; chromosome 10q24.

X Homo sapiens.

X WO200200677-A1.

X 03-JAN-2002.

X 07-JUN-2001; 2001WO-US018569.

XX  
 PR 07-JUN-2000; 2000US-0209467P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Birse CE, Rosen CA;  
 XX  
 XX WPI; 2002-147878/19.  
 DR N-PSDB; ABQ54872.  
 DR

XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
 PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian  
 PT cancer), immune disorders, cardiovascular disorders and neurological  
 PT diseases.  
 XX  
 PS Claim 11; SEQ ID NO 2927; 2922pp; English.

XX The invention relates to 2175 novel human ovarian antigens (ABP41054-  
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
 CC to the sequences of the invention. The invention additionally relates to  
 CC recombinant vectors and host cells comprising human ovarian antigen  
 CC polynucleotides, antibodies against human ovarian antigens, and the use  
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
 CC treating, prognosing or preventing various ovary and/or breast-related  
 CC disorders. Such conditions include ovarian cancer and breast cancer, and  
 CC metastatic tumours of ovarian or breast origin, reproductive system  
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine  
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis), and toxic  
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
 CC vaginitis), immune disorders (e.g., congenital and acquired  
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
 CC respiratory disorders, neurological disorders, gastrointestinal disorders  
 CC and urinary system disorders. Ovarian antigen polypeptides and  
 CC polynucleotides may also be used in screening for compounds which  
 CC modulate ovarian antigen expression or activity. The polynucleotides may  
 CC further be used for gene therapy, chromosome mapping, in the  
 CC identification of individuals and in forensic analysis, and the  
 CC polypeptides may be used as food additives or to prepare antibodies  
 CC useful in disease diagnosis, drug targeting and phenotyping. The present  
 CC sequence represents a human ovarian antigen of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 337 AA;

Query Match 100.0%; Score 510; DB 5; Length 337;  
 Best Local Similarity 100.0%; Pred. No. 7.6e-41;  
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGPRCLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60  
 Db 74 KTCYEGNGHFYRGKASTDTMGPRCLPWNATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 133  
 Qy 61 RRRPWCYVQVGLKPLVQECWVHDCADGK 88  
 Db 134 RRRPWCYVQVGLKPLVQECWVHDCADGK 161

RESULT 7

AAH68854

ID AAR68854 standard; protein; 365 AA.

XX AAR68854;

XX 16-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 22-NOV-1995 (first entry)

XX Delta 1-46 urokinase.

X: 13-JUN-2001; 2001WO-US018976.  
P: 20-JUN-2000; 2000US-0212874P.  
X: (UYPE-) UNIV PENNSYLVANIA.  
X: Cines DB, Higazi AA;  
X: WPI; 2002-122240/16.  
D: N-PSDB; AAD27078.  
X: Composition for modulating muscle cell and tissue contractility for  
P: treating atherosclerosis, asthma, hypertension, glaucoma, impotence,  
P: comprising domains from urokinase-type plasminogen activator.  
X: Claim 11; Fig 1D; 117pp; English.  
X: The invention relates to a composition comprising one or more domains of  
C: urokinase-type plasminogen activator (uPA). The composition is used to  
C: modulate the contractility and angiogenic activity of a mammalian muscle,  
C: endothelial cell or tissue. The composition is used for treating stroke,  
C: hypotension, hypertension, atherosclerosis, heart attack, microvascular  
C: occlusions, thrombotic microangiopathies, surgically induced thrombotic  
C: disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell  
C: invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,  
C: diabetic retinopathy, wound healing, clotting disorder, uterine  
C: contraction disorder, male impotence, respiratory disease or condition  
C: such as asthma, adult respiratory distress syndrome, primary pulmonary  
C: hypertension, microvascular thrombotic occlusion, and a disorder  
C: associated with chronic intrapulmonary fibrin formation. The present  
C: sequence is human urokinase-type plasminogen activator (uPA) amino  
C: terminal fragment (ATF)  
X: S: Sequence 135 AA;  
Query Match 100.0%; Score 510; DB 5; Length 135;  
Best Local Similarity 100.0%; Pred. No. 3.2e-41; Indels 0; Gaps 0;  
Matches 88; Conservative 0; Mismatches 0;  
Q: 1 KTCYEGNGHFGYRGKASTDTMGRCPLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNP DN 60  
D: 48 KTCYEGNGHFGYRGKASTDTMGRCPLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNP DN 107  
Q: 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
D: 108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135  
BLT 4  
AF 5549  
II AAE16549 standard; protein; 143 AA.  
X: AAE16549;  
X: 09-APR-2002 (first entry)  
X: Human uPA amino terminal fragment (ATF) and connecting peptide.  
X: Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;  
K: stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;  
X: microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;  
X: tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;  
K: clotting disorder; uterine contraction disorder; respiratory disease;  
K: adult respiratory distress syndrome; amino terminal fragment; ATF;  
K: male impotence.  
X: Homo sapiens.  
X: WO200197752-A2.  
P: 27-DEC-2001.

PF 13-JUN-2001; 2001WO-US018976.  
XX 20-JUN-2000; 2000US-0212874P.  
XX (UYPE-) UNIV PENNSYLVANIA.  
XX Cines DB, Higazi AA;  
XX WPI; 2002-122240/16.  
DR N-PSDB; AAD27082.  
XX Composition for modulating muscle cell and tissue contractility for  
PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,  
PT comprising domains from urokinase-type plasminogen activator.  
XX Claim 24; Fig 1H; 117pp; English.  
XX The invention relates to a composition comprising one or more domains of  
CC urokinase-type plasminogen activator (uPA). The composition is used to  
CC modulate the contractility and angiogenic activity of a mammalian muscle,  
CC endothelial cell or tissue. The composition is used for treating stroke,  
CC hypotension, hypertension, atherosclerosis, heart attack, microvascular  
CC occlusions, thrombotic microangiopathies, surgically induced thrombotic  
CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell  
CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,  
CC diabetic retinopathy, wound healing, clotting disorder, uterine  
CC contraction disorder, male impotence, respiratory disease or condition  
CC such as asthma, adult respiratory distress syndrome, primary pulmonary  
CC hypertension, microvascular thrombotic occlusion, and a disorder  
CC associated with chronic intrapulmonary fibrin formation. The present  
CC sequence is human urokinase-type plasminogen activator (uPA) amino  
CC terminal fragment (ATF) and connecting peptide  
XX SQ Sequence 143 AA;  
Query Match 100.0%; Score 510; DB 5; Length 143;  
Best Local Similarity 100.0%; Pred. No. 3.4e-41; Indels 0; Gaps 0;  
Matches 88; Conservative 0; Mismatches 0;  
Q: 1 KTCYEGNGHFGYRGKASTDTMGRCPLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNP DN 60  
D: 48 KTCYEGNGHFGYRGKASTDTMGRCPLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNP DN 107  
Q: 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88  
D: 108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135  
RESULT 5  
AAG75492  
ID AAG75492 standard; protein; 337 AA.  
XX AAG75492;  
XX 03-SEP-2001 (first entry)  
XX Human colon cancer antigen protein SEQ ID NO:6256.  
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;  
X: colorectal carcinoma; chromosome 10.  
XX Homo sapiens.  
XX WO200122920-A2.  
XX 05-APR-2001.  
XX 28-SEP-2000; 2000WO-US026524.  
XX 29-SEP-1999; 99US-0157137P.  
PR 03-NOV-1999; 99US-0163280P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA

X P P WO200197752-A2.  
X P P 27-DEC-2001.  
X P P 13-JUN-2001; 2001WO-US018976.  
X P P 20-JUN-2000; 2000US-0212874P.  
X P P (UYPE-) UNIV PENNSYLVANIA.  
X P P Cines DB, Higazi AA;  
X P P WPI; 2002-122240/16.  
X P P N-PSDB; AAD27075.  
X P P Composition for modulating muscle cell and tissue contractility for  
X P P treating atherosclerosis, asthma, hypertension, glaucoma, impotence,  
X P P comprising domains from urokinase-type plasminogen activator.  
X P P Claim 1; Fig 1A; 117pp; English.  
X P P The invention relates to a composition comprising one or more domains of  
X P P urokinase-type plasminogen activator (uPA). The composition is used to  
X P P modulate the contractility and angiogenic activity of a mammalian muscle,  
X P P endothelial cell or tissue. The composition is used for treating stroke,  
X P P hypotension, hypertension, atherosclerosis, heart attack, microvascular  
X P P occlusions, thrombotic microangiopathies, surgically induced thrombotic  
X P P disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell  
X P P invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,  
X P P diabetic retinopathy, wound healing, clotting disorder, uterine  
X P P contraction disorder, male impotence, respiratory distress syndrome, primary pulmonary  
X P P such as asthma, adult respiratory distress syndrome, and a disorder  
X P P hypertension; microvascular thrombotic occlusion, and a disorder  
X P P associated with chronic intrapulmonary fibrin formation. The present  
X P P sequence is human urokinase-type plasminogen activator (uPA) kringle  
X S Sequence 88 AA;  
Query Match 100.0%; Score 510; DB 5; Length 88;  
Best Local Similarity 100.0%; Pred. No. 2.2e-41;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Q 1 KTCYEGNGHFYRGKASTDTMGPRCLPWSATVLQOITYHAHRSDALQGLGKHNYCRNPDN 60  
D 1 KTCYEGNGHFYRGKASTDTMGPRCLPWSATVLQOITYHAHRSDALQGLGKHNYCRNPDN 60  
Q 61 RRPWCYVQVGLKPLVQECWHDCAQK 88  
D 61 RRPWCYVQVGLKPLVQECWHDCAQK 88  
K I A AAE16550 standard; protein; 96 AA.  
X X AAE16550;  
X X 09-APR-2002 (first entry)  
X X Human uPA kringle and connecting peptide.  
X X Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;  
X X stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;  
X X microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;  
X X tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;  
X X clotting disorder; uterine contraction disorder; respiratory disease;  
X X adult respiratory distress syndrome; male impotence.  
X X Homo sapiens.  
X X WO200197752-A2.  
X P 27-DEC-2001.

PD 27-DEC-2001.  
XX 13-JUN-2001; 2001WO-US018976.  
XX 20-JUN-2000; 2000US-0212874P.  
XX (UYPE-) UNIV PENNSYLVANIA.  
XX Cines DB, Higazi AA;  
XX WPI; 2002-122240/16.  
XX N-PSDB; AAD27083.  
XX Composition for modulating muscle cell and tissue contractility for  
XX treating atherosclerosis, asthma, hypertension, glaucoma, impotence,  
XX comprising domains from urokinase-type plasminogen activator.  
XX Claim 25; Fig 11; 117pp; English.  
XX The invention relates to a composition comprising one or more domains of  
XX urokinase-type plasminogen activator (uPA). The composition is used to  
XX modulate the contractility and angiogenic activity of a mammalian muscle,  
XX endothelial cell or tissue. The composition is used for treating stroke,  
XX hypotension, hypertension, atherosclerosis, heart attack, microvascular  
XX occlusions, thrombotic microangiopathies, surgically induced thrombotic  
XX disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell  
XX invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,  
XX diabetic retinopathy, wound healing, clotting disorder, uterine  
XX contraction disorder, male impotence, respiratory distress syndrome, primary pulmonary  
XX such as asthma, adult respiratory distress syndrome, and a disorder  
XX hypertension; microvascular thrombotic occlusion, and a disorder  
XX associated with chronic intrapulmonary fibrin formation. The present  
XX sequence is human urokinase-type plasminogen activator (uPA) kringle and  
XX connecting peptide  
XX Sequence 96 AA;  
Query Match 100.0%; Score 510; DB 5; Length 96;  
Best Local Similarity 100.0%; Pred. No. 2.3e-41;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWSATVLQOITYHAHRSDALQGLGKHNYCRNPDN 60  
DB 1 KTCYEGNGHFYRGKASTDTMGPRCLPWSATVLQOITYHAHRSDALQGLGKHNYCRNPDN 60  
QY 61 RRPWCYVQVGLKPLVQECWHDCAQK 88  
DB 61 RRPWCYVQVGLKPLVQECWHDCAQK 88  
RESULT 3  
AAE16545  
ID AAE16545 standard; protein; 135 AA.  
XX AAE16545;  
XX AC AAE16545;  
XX 09-APR-2002 (first entry)  
XX Human urokinase-type plasminogen activator amino terminal fragment (ATP).  
XX Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;  
XX stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;  
XX microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;  
XX tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;  
XX clotting disorder; uterine contraction disorder; respiratory disease;  
XX adult respiratory distress syndrome; amino terminal fragment; ATP;  
XX male impotence.  
XX Homo sapiens.  
XX WO200197752-A2.  
XX 27-DEC-2001.

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

Of protein - protein search, using sw model

Run on: May 25, 2004, 14:43:00 ; Search time 21.518 Seconds  
(without alignments)  
1155.508 Million cell updates/sec

T: e: US-09-880-503-1  
P: ect score: 510  
S: ence: 1 KTCYEGNGHFYCKASTDM.....QVGLKPLVQECMVHDCADGK 88

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Sequenced: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

```

M:      num DB seq length: 0
M:      num DB seq length: 2000000000
Pc      -processing: Minimum Match 0%
           Maximum Match 10%
           Listing first 75

```

```
Dz : base :
1_ Genesep29Jan04.*
1: Genesep21960s.*
2: Genesep1950s.*
3: Genesep22008s.*
4: Genesep2001s.*
5: Genesep2002s.*
6: Genesep2003aa.*
7: Genesep2003ba.*
8: Genesep2004s.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Re Id.	Query			DB	ID	Description
	Score	Match	Length			
1	510	100.0	88	5	AAE16542	AAe16542 Human uro
2	510	100.0	96	5	AAE16550	AAe16550 Human uro
3	510	100.0	135	5	AAE16545	AAe16545 Human uro
4	510	100.0	143	5	AAE16549	AAe16549 Human uro
5	510	100.0	337	4	AAg75492	AAg75492 Human col
6	510	100.0	337	5	ABP41795	ABP41795 Human ova
7	510	100.0	365	2	AAr68854	AAr68854 Delta 1-4
8	510	100.0	378	2	AAW13635	AAW13635 Human pro
9	510	100.0	386	2	AAr66366	AAr66366 Bifunctio
10	510	100.0	389	2	AAW13636	AAW13636 Human pro
11	510	100.0	390	2	AAr66245	AAr66245 Bifunctio
12	510	100.0	390	2	AAr66247	AAr66247 Bifunctio
13	510	100.0	392	2	AAr66260	AAr66260 Bifunctio
14	510	100.0	392	2	AAr66264	AAr66264 Bifunctio
15	510	100.0	392	2	AAr66255	AAr66255 Bifunctio
16	510	100.0	392	2	AAr66259	AAr66259 Bifunctio
17	510	100.0	392	2	AAr66261	AAr66261 Bifunctio
18	510	100.0	392	2	AAr66258	AAr66258 Bifunctio
19	510	100.0	392	2	AAr66263	AAr66263 Bifunctio
20	510	100.0	392	2	AAr66254	AAr66254 Bifunctio
21	510	100.0	392	2	AAr66256	AAr66256 Bifunctio
22	510	100.0	392	2	AAr66257	AAr66257 Bifunctio
23	510	100.0	393	2	AAr66251	AAr66251 Bifunctio
24	510	100.0	393	2	AAr66253	AAr66253 Bifunctio
25	510	100.0	393	2	AAr66249	AAr66249 Bifunctio

## ALIGNMENTS

RESULT 1  
AAE16542  
ID AAE16542 standard; protein; 88 AA.

AA	AAE16542;	
AC		
XX		
DT	09-APR-2002	(first entry)

Human urokinase-type plasminogen activator (uPA) kringle.

Human, urokinase-type plasminogen activator; uPA; therapy; hypertension; stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder; microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma; tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing; clotting disorder; uterine contraction disorder; respiratory disease; male impotence; adult respiratory distress syndrome.

Homo sapiens.



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2004, 14:47:10 ; Search time 7.80994 Seconds  
(without alignments)  
1662.947 Million cell updates/sec

Title: US-09-880-503-4  
Perfect score: 793  
Sequence: 1 SNEHQVPSNCDLNGGTCV.....QVGLKPLVQECWVHDCADGK 135

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : PIR 78:\*\*

1: PIR1:\*\*  
2: PIR2:\*\*  
3: PIR3:\*\*  
4: PIR4:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	793	100.0	431	1	UKHU	u-plasminogen acti
2	713.5	90.0	431	1	UKBY	u-plasminogen acti
3	608	76.7	442	1	UKPG	u-plasminogen acti
4	579	73.0	433	1	JN0560	u-plasminogen acti
5	571	72.0	432	1	S18932	u-plasminogen acti
6	548	69.1	433	1	UKMS	u-plasminogen acti
7	335.5	42.3	477	1	A34369	u-plasminogen acti
8	335.5	42.3	477	2	JS0598	u-plasminogen acti
9	334.5	42.2	434	1	A35005	u-plasminogen acti
10	328.5	41.4	211	2	I38088	t-plasminogen acti
11	328.5	41.4	431	2	JS0599	t-plasminogen acti
12	328.5	41.4	562	1	UKHUT	t-plasminogen acti
13	321.5	40.5	559	1	A35029	t-plasminogen acti
14	315.5	39.8	559	1	A29941	t-plasminogen acti
15	310.5	39.2	477	2	JS0597	t-plasminogen acti
16	264.5	33.4	558	2	JS0578	plasma hyaluronan-
17	260.5	32.8	685	1	A46588	hepatocyte growth
18	257.5	32.5	560	1	JC4795	plasma hyaluronan-
19	257	32.4	603	2	S28941	coagulation factor
20	233	29.4	615	1	KFHU12	coagulation factor
21	223	28.1	394	2	JS0600	t-plasminogen acti
22	217	27.4	593	2	S45281	coagulation factor
23	169	21.3	685	1	A48289	neutrotrophic recep
24	162.5	20.5	4548	1	S00857	apoptosis(a) (EC
25	161	20.3	1420	2	A32869	apoptosis(a)
26	160	20.2	123	2	C61545	plasmin (EC 3.4.21
27	159	20.1	120	2	E61545	plasmin (EC 3.4.21
28	158.5	20.0	937	2	A45082	neutrotrophic recep
29	157.5	19.9	460	2	B61545	plasmin (EC 3.4.21

#### RESULT 1

##### UKHU

u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human  
N/Alternate names: cellular plasminogen activator; urokinase; urokinase-type p  
N/Contains: urokinase-type plasminogen activator chain A; urokinase-type plasm  
in form  
C/Species: Homo sapiens (man)  
C/Date: 17-Dec-1982 #sequence\_revision 04-Dec-1986 #text change 15-Sep-2000  
C/Accession: A00931; I52209; JTO102; A37561; I38102; S65783; A37562; A  
R/Riccio, A.; Grimaldi, G.; Verde, P.; Sebastio, G.; Boast, S.; Blas, F.  
Nucleic Acids Res. 13, 2759-2771, 1985  
A/Title: The human urokinase-plasminogen activator gene and its promoter.  
A/Reference number: A00931; MUID:85215647; PMID:2987867  
A/Accession: A00931  
A/Molecule type: DNA  
A/Residues: 1-431 <RIC>  
A/Cross-references: GB:X02419; NID:937601; PIDN:CAA26268.1; PID:91834524  
A/Note: the authors translated the codon ATG for residue 214 as Ile  
R/Nagamine, Y.; Pearson, D.; Grattan, M.  
Biochem. Biophys. Res. Commun. 132, 563-569, 1985  
A/Title: Exon-intron boundary sliding in the generation of two mRNAs coding fo  
A/Reference number: I52209; MUID:86050639; PMID:3933505

#### ALIGNMENTS

30	156	19.7	812	1	PLBO	plasmin (EC 3.4.21
31	155	19.5	790	1	PLFG	plasmin (EC 3.4.21
32	154	19.4	89	2	A60140	plasmin (EC 3.4.21
33	154	19.4	810	1	PLHU	plasmin (EC 3.4.21
34	153.5	19.4	810	2	I46260	plasmin (EC 3.4.21
35	152.5	19.2	810	2	B30848	plasmin (EC 3.4.21
36	152	19.2	169	2	A48522	plasmin (EC 3.4.21
37	148	18.7	943	2	B45082	neutrotrophic recep
38	147	18.5	812	1	PLMS	plasmin (EC 3.4.21
39	144.5	18.2	806	2	T18840	hypothetical prote
40	144	18.2	711	1	A47136	macrophage-stimula
41	141.5	17.8	455	2	A61545	plasmin (EC 3.4.21
42	141.5	17.8	728	1	JH0579	hepatocyte growth
43	141	17.8	411	2	I51285	hepatocyte growth
44	141	17.8	710	1	I51283	hepatocyte growth
45	139	17.5	716	1	A40332	macrophage-stimula
46	139	17.5	728	1	A35644	hepatocyte growth
47	139	17.5	728	1	A60185	hepatocyte growth
48	138	17.4	716	1	JC5061	macrophage-stimula
49	136	17.2	2869	2	T18518	apolipoprotein(a)
50	136	17.2	336	2	S33879	plasmin precursor
51	125.5	15.8	618	2	A35827	thrombin (EC 3.4.2
52	124	15.6	622	1	TBHU	thrombin (EC 3.4.2
53	122	15.4	946	1	A47299	ror-related recep
54	117.5	14.8	2352	2	T30201	Notch homolog prot
55	117	14.8	2524	2	A35844	Xotch protein - Af
56	115	14.5	625	1	TBBO	thrombin (EC 3.4.2
57	113	14.2	617	2	S10511	thrombin (EC 3.4.2
58	108.5	13.7	1620	2	T27283	hypothetical prote
59	106	13.4	2531	2	T31070	notch homolog - se
60	105.5	13.3	2109	1	I50421	aggregran precursor
61	103	13.0	1203	2	A49175	Notch B protein -
62	101.5	12.8	2555	2	A40043	notch protein homo
63	101.5	12.8	2703	1	A24420	notch protein - fr
64	101	12.7	379	2	A59180	Wnt inhibitory fac
65	101	12.7	2471	2	A49128	cell-face determin
66	101	12.7	2531	2	S18188	notch protein homo
67	100.5	12.7	1523	2	T13953	MEGFS protein - ra
68	100.5	12.7	2139	2	A35672	crumbs protein - f
69	100	12.6	387	2	B49175	Notch A protein -
70	100	12.6	2531	2	A46019	notch-1 protein -
71	99.5	12.5	861	2	A48825	Notch homolog Motc
72	99.5	12.5	2437	2	S42812	transmembrane prot
73	99	12.5	2809	2	T30213	G-cadherin - sea u
74	99	12.5	3034	2	T14119	Seven-pass transme
75	98.5	12.4	686	2	JC7569	Delta-4 protein -

A/Accession: I52209  
A/Status: preliminary; translated from GB/EMBL/DBDJB  
A/Molecule type: DNA  
A/Residues: 145-161 <NAG1>  
A/Cross-references: GB:K03027; NID:G340174; PIDN:AAA61257.1; PID:G340175  
R:Nagai, M.; Hiramatsu, R.; Kaneda, T.; Hayasuke, N.; Arimura, H.; Nishida, M.; Suyama, Gene 36, 183-188, 1985  
A/Title: Molecular cloning of cDNA coding for human preprourokinase.  
A/Reference number: JTO102; MUID:86056954; PMID:2415429  
A/Accession: JTO102  
A/Molecule type: mRNA  
A/Residues: 1-213, 'I', 215-431 <NAG2>  
A/Cross-references: GB:K03226; NID:G340155; PIDN:AA097138.1; PID:G340158; GB:D00244; NID:G340159  
R:Verde, P.; Stoppelli, M.P.; Galeffi, P.; Di Nocera, P.; Blasi, F.  
Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984  
A/Title: Identification and primary sequence of an unspliced human urokinase poly(A)+ RNA  
A/Reference number: A37561; MUID:84272706; PMID:6589620  
A/Accession: A37561  
A/Molecule type: mRNA  
A/Residues: 66-431 <VER>  
A/Cross-references: GB:D00244; NID:G220138  
R:Jacobs, P.; Cravador, A.; Loriau, R.; Brockly, F.; Colau, B.; Chuchana, P.; van Elsen, DNA 4, 139-146, 1985  
A/Title: Molecular cloning, sequencing, and expression in *Escherichia coli* of human prepro-urokinase cDNA  
A/Reference number: I38102; MUID:85203359; PMID:3888571  
A/Accession: I38102  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-150, 'W', 152-213, 'I', 215-385, 'C', 387-429, 'V', 431 <JAC>  
A/Cross-references: EMBL:X02760; NID:G335297; PIDN:CAA26335.1; PID:G335298  
R:Yoshimoto, M.; Ushiyama, Y.; Sakai, M.; Tamaki, S.; Hara, H.; Takahashi, K.; Sawasaki, Biochim. Biophys. Acta 1293, 83-89, 1996  
A/Title: Characterization of single chain urokinase-type plasminogen activator with a no  
A/Reference number: S65783; MUID:96186279; PMID:8652631  
A/Accession: S65783  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 21-140, 'L', 142-213, 'I', 215-431 <YOS>  
A/Cross-references: EMBL:D11143; NID:G1311467; PIDN:BA01919.1; PID:gl199928  
R:Gunzler, W.A.; Steffens, G.J.; Otting, F.; Kim, S.M.A.; Frankus, E.; Fliche, L.  
Hoppe-Seyler's Z. Physiol. Chem. 363, 1155-1165, 1982  
A/Title: The primary structure of high molecular mass urokinase from human urine.  
A/Reference number: A37562; MUID:83055084; PMID:6754569  
A/Accession: A37562  
A/Molecule type: protein  
A/Residues: 21-177 <GUN>  
R:Schaller, J.; Nick, H.; Rickli, E.E.; Gillesse, D.; Lergier, W.; Studer, R.O.  
Eur. J. Biochem. 125, 251-257, 1982  
A/Title: Human low-molecular-weight urinary urokinase. Partial characterization and preli  
A/Reference number: A37563; MUID:93003608; PMID:6749491  
A/Accession: A37563  
A/Molecule type: protein  
A/Residues: 156-176; 179-193, 'T', 195, 'T', 197-224 <SCH>  
R:Steffens, G.J.; Gunzler, W.A.; Otting, F.; Frankus, E.; Fliche, L.  
Hoppe-Seyler's Z. Physiol. Chem. 363, 1043-1058, 1982  
A/Title: The complete amino acid sequence of low molecular mass urokinase from human uri  
A/Reference number: A37564; MUID:83055099; PMID:6754572  
A/Accession: A37564  
A/Molecule type: protein  
A/Residues: 158-410 <STE>  
R:Kantzer, E.J.; Buko, A.; Menon, G.; Saxin, V.K.  
Biochem. Biophys. Res. Commun. 171, 401-406, 1990  
A/Title: Carbohydrate composition and presence of a fucose-protein linkage in recombinan  
A/Reference number: A35689; MUID:90365737; PMID:2393398  
A/Accession: A35689  
A/Molecule type: protein  
A/Residues: 21-30, 'X', 32, 'X', 34-38, 'X', 40-43 <KEN>  
A/Note: Identification of a fucose and attempt to determine its attachment site  
R:Rabbani, S.A.; Desjardins, J.; Bell, A.W.; Banville, D.; Mazar, A.; Henkin, J.; Goltzm  
Biochem. Biophys. Res. Commun. 173, 1058-1064, 1990  
A/Title: An amino-terminal fragment of urokinase isolated from a prostate cancer cell li  
A/Reference number: A36697; MUID:91097529; PMID:2125213  
A/Accession: A36697

A/Molecule type: protein  
A/Residues: 21-34 <RAB>  
R:Li, X.; Bokman, A.M.; Llinas, M.; Smith, R.A.G.; Dobson, C.M.  
submitted to the Brookhaven Protein Data Bank, July 1993  
A/Reference number: A51255; PDB:1KDU  
A/Contents: annotation; conformation and disulfide bond assignments by (1)H-NM  
R:Li, X.; Smith, R.A.G.; Dobson, C.M.  
Biochemistry 31, 9562-9571, 1992  
A/Title: Sequential (1)H NMR assignments and secondary structure of the kringl  
A/Reference number: A44375; MUID:93003110; PMID:1327118  
A/Contents: annotation; conformation and disulfide bond assignments by (1)H-NM  
R:Hansen, A.P.; Petros, A.M.; Meadows, R.P.; Nettesheim, D.G.; Mazar, A.P.; Ol  
submitted to the Brookhaven Protein Data Bank, January 1994  
A/Reference number: A66822; PDB:1URK  
A/Contents: annotation; conformation and disulfide bond assignments by (1)H-NM  
R:Spraggon, G.S.; Phillips, C.; Nowak, U.K.; Ponting, C.P.; Saunders, D.; Dobs  
submitted to the Brookhaven Protein Data Bank, July 1995  
A/Reference number: A66058; PDB:1LMW  
A/Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 168-175  
C/Comment: This enzyme is found in urine in a high molecular mass form, consis  
C/Comment: Urokinase-type plasminogen activator proteolytically activates plas  
C/Genetics:  
A/Gene: GDB:PLAU  
A/Cross-references: GDB:119497; OMIM:191840  
A/Map position: 10q24-10q24  
A/Introns: 19/3; 29/1; 65/1; 123/2; 154/1; 227/2; 277/1; 324/1; 373/3  
C/Function:  
A/Description: proteolytically activates plasminogen  
A/Pathway: fibrinolysis  
C/Superfamily: urokinase-type plasminogen activator; EGF homology; kringle hom  
C/Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serin  
F:1-20/Domain: signal sequence #status predicted <SIG>  
F:21-431/Product: urokinase-type plasminogen activator, single chain form #sta  
F:21-177/Product: urokinase-type plasminogen activator chain A #status experim  
F:31-62/Domain: EGF homology <EGF>  
F:70-151/Domain: kringle homology <KR>  
F:156-177/Product: urokinase-type plasminogen activator chain B #status exper  
F:178-419/Domain: trypsin homology <TRY>  
F:31-39,33-51,53-62,70-151,91-133,123-146,168-299,209-225,217-288,313-382,345-  
F:38/Binding site: carbohydrate (Thr) (covalent) #status predicted  
F:178-179/cleavage site: Lys-ile (plasmin) #status experimental  
F:224,275,376/Active site: His, Asp, Ser #status experimental  
F:322/Binding site: carbohydrate (Asn) (covalent) #status experimental  
Query Match 100.0%; Score 793; DB 1; Length 431;  
Best Local Similarity 100.0%; Pred. NO. 3.3e-64; Mismatches 0; Gaps 0  
Matches 135; Conservative 0; Indels 0  
QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHPIRG 60  
Db 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHPIRG 80  
QY 61 KASDTDMGRPCLPWNSATVLQOQTHAHRSDALQLGLGKHCYCRNPDRRPPWCYVQGLK 120  
Db 81 KASDTDMGRPCLPWNSATVLQOQTHAHRSDALQLGLGKHCYCRNPDRRPPWCYVQGLK 140  
QY 121 FLVQSCMVHDCADGK 135  
Db 141 FLVQSCMVHDCADGK 155  
RESULT 2  
UKBAY  
C/Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)  
C/Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 18-Jun-1999  
C/Accession: S14687; S08651  
R:Au, Y.P.T.; Wang, T.W.; Clowes, A.W.  
Nucleic Acids Res. 18, 3411, 1990  
A/Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type  
A/Reference number: S14687; MUID:90287734; PMID:2113276  
A/Accession: S14687

A;Molecule type: mRNA  
A;Residues: 1-433 <AUV>  
A;Cross-references: EMBL:X51935; NID:G38130; PIDN:CAA3200.1; PID:G38131  
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try  
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase  
F;1-20/Domain: signal sequence #status predicted <SIG>  
F;21-176/Product: plasminogen activator chain A #status predicted <ACH>  
F;30-61/Domain: EGF homology <EGF>  
F;69-150/Domain: kringle homology <KRG>  
F;178-433/Product: plasminogen activator chain B #status predicted <BCH>  
F;178-421/Domain: trypsin homology <TRY>  
F;167-298,208-224,216-287,315-384,347-363,374-402/Disulfide bonds: #status predicted  
F;223,274,378/Active site: His, Asp, Ser #status predicted  
F;324/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 90.0%; Score 713.5; DB 1; Length 433;  
Best Local Similarity 91.9%; Pred. No. 4,9e-57;  
Matches 124; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 SNEHQVPSNCDLNGGTCVSNKYFSNIHWNCNCPKKGQHQCEIDKSKTCYEGNGHFYRG 60  
DB 21 SREL-QVPSDCGLNGGTCVSNKYFSNIHWNCNCPKKGQHQCEIDKSKTCYEGNGHFYRG 79

QY 61 KASTDTMGRCPLPWNATVLOQTYHAHRS DALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120  
DB 80 KASTDTMGRCPLPWNATVLOQTYHAHRS DALQLGLGKHNYCRPNDRRRPWCYVQVGLK 139

QY 121 PLVQECMVHDCADGK 135  
DB 140 QRQVQECMVHDCADGK 154

RESULT 3  
UKFG  
u-plasminogen activator (EC 3.4.21.73) precursor - pig  
N;Alternate names: uPA  
C;Species: Sus scrofa domestica (domestic pig)  
C;Date: 04-Dec-1986 #sequence\_revision 17-Mar-1987 #text\_change 07-Aug-1998  
C;Accession: A00932

R;Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, E.  
Nucleic Acids Res. 12, 9525-9541, 1984  
A;Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.  
A;Reference number: A00932; MUID:85087954; PMID:6096832  
A;Accession: A00932  
A;Molecule type: DNA  
A;Residues: 1-240, 'H', 242-442 <NAG1>  
A;Experimental source: kidney cell line LLC-PK1  
R;Nagamine, Y.  
Submitted to the Protein Sequence Database, December 1986  
A;Reference number: A37566  
A;Contents: annotation; correction to residue 241  
C;Genetics:  
A;Introns: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3  
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try  
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase  
F;1-20/Domain: signal sequence #status predicted <SIG>  
F;21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>  
F;33-64/Domain: EGF homology <EGF>  
F;72-153/Domain: kringle homology <KRG>  
F;190-442/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>  
F;190-430/Domain: trypsin homology <TRY>  
F;152/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;179-310,220-236,228-239,324-393,356-372,383-411/Disulfide bonds: #status predicted  
F;235,286,387/Active site: His, Asp, Ser #status predicted

Query Match 76.7%; Score 608; DB 1; Length 442;  
Best Local Similarity 78.4%; Pred. No. 1.6e-47;  
Matches 107; Conservative 11; Mismatches 17; Indels 2; Gaps 1;

QY 1 SNEHQV--PSNCDLNGGTCVSNKYFSNIHWNCNCPKKGQHQCEIDKSKTCYEGNGHFY 58  
DB 21 SHELQESGASNGCLNGGCKVYKFSNIQRSCPKFKQGEHCIEDTTSQTCFEGNGHGY 80

QY 59 RGKASTDTMGRCPLPWNATVLOQTYHAHRS DALQLGLGKHNYCRPNDRRRPWCYVQV 118  
DB 81 RGKANTNTGRCPLPWNATVLLNTVHAHSPDALQLGLGKHNYCRPNDRRRPWCYVQV 140

QY 119 LKPLVQECMVHDCADGK 135  
DB 141 LKQLVQECMVHDCADGK 157

RESULT 4  
JN0560  
u-plasminogen activator (EC 3.4.21.73) precursor - bovine  
N;Alternate names: uPA  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 16-Jul-1999  
C;Accession: JN0560  
R;Kaezschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schleuning, W.D.  
Gene 125, 177-183, 1993  
A;Title: Bovine urokinase-type plasminogen activator and its receptor: cloning and inc  
A;Reference number: JN0560; MUID:93216119; PMID:8385052  
A;Accession: JN0560  
A;Molecule type: mRNA  
A;Residues: 1-433 <KRA>  
A;Cross-references: GB:I03546; NID:G16380; PIDN:AAA51419.1; PID:G163801  
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle hom  
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase  
F;1-20/Domain: signal sequence #status predicted <SIG>  
F;21-179/Product: plasminogen activator chain A #status predicted <MA1>  
F;21-179/Product: urokinase-type plasminogen activator chain A #status predict  
F;33-64/Domain: EGF homology <EGF>  
F;72-153/Domain: kringle homology <KRG>  
F;181-433/Product: plasminogen activator chain B #status predicted <MA2>  
F;181-421/Domain: trypsin homology <TRY>  
F;170-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status pre  
F;226,277,378/Active site: His, Asp, Ser #status predicted

Query Match 73.0%; Score 579; DB 1; Length 433;  
Best Local Similarity 73.0%; Pred. No. 6.7e-45;  
Matches 100; Conservative 13; Mismatches 22; Indels 2; Gaps 1

QY 1 SNEHQV--PSNCDLNGGTCVSNKYFSNIHWNCNCPKKGQHQCEIDKSKTCYEGNGHFY 58  
DB 21 SNEVHKESGSGNCGCLNGGCKVYKFSNIQRSCPKFKQGEHCIEDTSTCYQGNHGY 80

QY 59 RGKASTDTMGRCPLPWNATVLOQTYHAHRS DALQLGLGKHNYCRPNDRRRPWCYVQV 118  
DB 81 RGKANDLSDRCPLANDSPVLLKMYHAHRS DALQLGLGKHNYCRPNDRRRPWCYVQI 140

QY 119 LKPLVQECMVHDCADGK 135  
DB 141 LKQFVQFCMVQDCSVGK 157

RESULT 5  
S18932  
u-plasminogen activator (EC 3.4.21.73) precursor - rat  
N;Alternate names: plasminogen activator, urokinase-type; urinary plasminogen  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 18-Oct-1989 #sequence\_revision 10-Feb-1995 #text\_change 18-Jun-1999  
C;Accession: S24604; I60186; I53472; S18932  
R;Rabbani, S.A.  
Submitted to the EMBL Data Library, April 1992  
A;Reference number: S24604  
A;Accession: S24604  
A;Molecule type: mRNA  
A;Residues: 1-15, 'H', 17-23, 'G', 25-331, 'N', 333-432 <RAB>  
A;Cross-references: EMBL:X65651; NID:G57456; PIDN:CAA46601.1; PID:G57457  
A;Experimental source: tissue kidney  
R;Henderson, B.R.; Tansey, W.P.; Phillips, S.M.; Ramshaw, I.A.; Kefferd, R.F.  
Cancer Res. 52, 2489-2496, 1992  
A;Title: Transcriptional and posttranscriptional activation of urokinase plasm  
A;Reference number: I60186; MUID:92233409; PMID:1568219  
A;Accession: I60186

A;Status: preliminary; translated from GB/EMBL/DBDJ  
 A;Molecule type: mRNA  
 A;Residues: 1-432 <RES>  
 A;Cross-references: EMBL:X63434; NID:G57465; PIDN:CAA45028.1; PID:G57466  
 A;Experimental source: EMBL:Strain Fischer.344; tissue mammary  
 R;Ragno, P.; Casano, S.; Degen, J.; Kessler, C.; Blasi, F.; Rossi, G.  
 FEBS Lett. 306, 193-198, 1992  
 A;Title: The receptor for the plasminogen activator of urokinase type is up-regulated in  
 A;Reference number: I53472; MUID:92339549; PMID:1321734  
 A;Accession: I53472  
 A;Status: preliminary; translated from GB/EMBL/DBDJ  
 A;Molecule type: DNA  
 A;Residues: 31-62 <RES2>  
 A;Cross-references: EMBL:X66907; NID:G396200; PIDN:CAA47356.1; PID:G938279  
 C;Genetics: uba  
 C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try  
 C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase  
 F;1-19/Domain: signal sequence #status predicted <SIG>  
 F;20-177/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>  
 F;31-62/Domain: EGF homology <EGF>  
 F;70-151/Domain: kringle homology <KRG>  
 F;179-432/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>  
 F;179-420/Domain: trypsin homology <TRY>  
 F;168-300.210-226.218-289.314-383.346-362.373-401/Disulfide bonds: #status predicted  
 F;225.276.377/Active site: His, Asp, Ser #status predicted  
 Query Match: 72.0%; Score 571; DB 1; Length 432;  
 Best Local Similarity 77.2%; Pred. No. 3.5e-44;  
 Matches 98; Conservative 8; Mismatches 21; Indels 0; Gaps 0;  
 QY 9 SNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRGKASTDTMG 68  
 DB 29 SNGCCQGGVCVSYKYFSIRRCSCPKFKGHCIDTSTKYVHGNGQYRGKANTDYG 88  
 QY 69 RPLCPWNSATVLOQTHAHRSDALQLGLGKHNYCRPNDRRPPWCYVQGLKPLVQECMV 128  
 DB 89 RPLCLANSPAVLQOQTHAHRSDALSLGLGKHNYCRPNDRRPPWCYVQGLKPLVQECMV 148  
 QY 129 HDCADGK 135  
 DB 149 QDCLSLK 155  
 RESULT 6  
 UKMS  
 u-plasminogen activator (EC 3.4.21.73) precursor - mouse  
 C;Species: Mus musculus (house mouse)  
 C;Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 18-Jun-1999  
 C;Accession: A29420; A24615  
 R;Degen, S.J.F.; Heckel, J.L.; Reich, E.; Degen, J.L.  
 Biochemistry 26, 8270-8279, 1987  
 A;Title: The murine urokinase-type plasminogen activator gene.  
 A;Reference number: A29420; MUID:88163489; PMID:2831940  
 A;Accession: A29420  
 A;Molecule type: DNA  
 A;Residues: 1-433 <DEG>  
 A;Cross-references: GB:M17922; NID:G202296; PIDN:AAA40539.1; PID:G202297  
 R;Belin, D.; Vassalli, J.D.; Combepine, C.; Godeau, F.; Nagamine, Y.; Reich, E.; Kocher,  
 Eur. J. Biochem. 148, 225-232, 1985  
 A;Title: Cloning, nucleotide sequencing and expression of cDNAs encoding mouse urokinase  
 A;Reference number: A24615; MUID:85179474; PMID:2985383  
 A;Accession: A24615  
 A;Molecule type: mRNA  
 A;Residues: 1-433 <BEL>  
 A;Cross-references: GB:X02389; NID:G55127; PIDN:CAA26231.1; PID:G55128  
 C;Genetics:  
 A;Introns: 19/3; 30/1; 66/1; 124/2; 155/1; 229/2; 279/1; 326/1; 375/3  
 C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try  
 C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase  
 F;1-20/Domain: signal sequence #status predicted <SIG>  
 F;21-178/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>  
 F;32-63/Domain: EGF homology <EGF>

F;71-152/Domain: kringle homology <KRG>  
 F;180-433/Product: urokinase-type plasminogen activator chain B #status predic  
 F;180-421/Domain: trypsin homology <TRY>  
 F;169-301.211-227.219-290.315-384.347-402/Disulfide bonds: #status pre  
 F;226.277.378/Active site: His, Asp, Ser #status predicted  
 Query Match: 59.1%; Score 548; DB 1; Length 433;  
 Best Local Similarity 71.7%; Pred. No. 4.2e-42;  
 Matches 91; Conservative 13; Mismatches 23; Indels 0; Gaps 0  
 QY 9 SNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRGKASTDTMG 68  
 DB 30 SNGCCQGGVCVSYKYFSIRRCSCPKFKGHCIDASKTCVHGNGDSYRGKANTDTGK 89  
 QY 69 RPLCPWNSATVLOQTHAHRSDALQLGLGKHNYCRPNDRRPPWCYVQGLKPLVQECMV 128  
 DB 90 RPLCLANSPAVLQOQTHAHRSDALSLGLGKHNYCRPNDRRPPWCYVQGLKPLVQECMV 149  
 QY 129 HDCADGK 135  
 DB 150 HDCLSLK 156  
 RESULT 7  
 A34369  
 t-plasminogen activator (EC 3.4.21.68) precursor - false vampire bat (Megaderm  
 C;Species: Megaderma lyra  
 C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C;Accession: A34369  
 R;Gadell, S.J.; Duong, L.T.; Diehl, R.E.; York, J.D.; Hare, T.R.; Register, R  
 J. Biol. Chem. 264, 17947-17952, 1989  
 A;Title: Isolation, characterization, and cDNA cloning of a vampire bat saliva  
 A;Reference number: A34369; MUID:90036867; PMID:2509450  
 A;Accession: A34369  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-477 <GAR>  
 A;Cross-references: GB:J05082; NID:G166080; PIDN:AAA31596.1; PID:G166081  
 C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I  
 C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
 F;1-21/Domain: signal sequence #status predicted <SIG>  
 F;22-36/Domain: propeptide #status predicted <PRO>  
 F;37-477/Product: plasminogen activator #status predicted <PLA>  
 F;42-79/Domain: fibronectin type I repeat homology <1FA>  
 F;87-120/Domain: EGF homology <EGF>  
 F;128-209/Domain: kringle homology <KRG>  
 F;226-471/Domain: trypsin homology <TRY>  
 F;42-72.70-79.87-98.92-109.111-120.128-209.149-191.180-204.214-345.257-273.265  
 F;272.321.428/Active site: His, Asp, Ser #status predicted  
 Query Match: 42.3%; Score 335.5; DB 1; Length 477;  
 Best Local Similarity 46.3%; Pred. No. 6.8e-23;  
 Matches 62; Conservative 17; Mismatches 50; Indels 5; Gaps 2  
 QY 3 ELHVP----SNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHY 58  
 DB 78 QCHTVPKSCSELRCFNGGTCQAAASDF--VCQCPKGYTGKQCEVTHATCYKQGVTV 136  
 QY 59 RGKASTDTMGPRCPFWNSATVLOQTHAHRSDALQLGLGKHNYCRPNDRRPPWCYVQGL 118  
 DB 137 RGTWSTSSGACQINWNSNLLTRTYRGRSDAITLGLGHNHYCRPNDRRPPWCYVQGL 196  
 QY 119 LKPLVQECMVHDC 132  
 DB 197 SKFILEFCVPCVS 210  
 RESULT 8  
 JS0598  
 t-plasminogen activator (EC 3.4.21.68) alpha-2 precursor - common vampire bat  
 N;Alternate names: tissue plasminogen activator  
 C;Species: Desmodus rotundus (common vampire bat)  
 C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999

C;Accession: J050598  
R;Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Dor  
Gene 105, 229-237, 1991  
A;Title: The plasminogen activator family from the salivary gland of the vampire bat Des  
A;Reference number: J050597; MUID:92039036; PMID:1937019  
A;Accession: J050598  
A;Molecule type: mRNA  
A;Residues: 1-477 <KRA>  
A;Cross-references: GB:M63988; NID:g166074; PIDN:AAA31593.1; PID:g166075  
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom  
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;122-36/Domain: propeptide #status predicted <PRO>  
F;137-477/Product: plasminogen activator alpha-2 #status predicted <PLA>  
F;142-79/Domain: fibronectin type I repeat homology <EFG>  
F;187-120/Domain: EGF homology <EGF>  
F;128-209/Domain: kringle homology <KRG>  
F;226-471/Domain: trypsin homology <TRY>  
F;422-72/70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4  
F;185,398/Binding site: Carbohydrate (Asn) (covalent) #status predicted  
F;185-226/Cleavage site: His-Ser (plamin) #status predicted  
F;272,321,428/Active site: His, Asp, Ser #status predicted  
Query Match 42.3%; Score 335.5; DB 2; Length 477;  
Best Local Similarity 46.3%; Pred. No. 6,8e-23;  
Matches 62; Conservative 17; Mismatches 50; Indels 5; Gaps 2;  
QY 3 ELHQVP-----SNCDLNGTGVSNKYFNIHWCNPKKFGQHCHEIDKSKTCYEGNGHFY 58  
DB 78 QCHTVPVKSCSELRCFNGTQWQAASFSDP-VCCQPKGYTKQCEVDTHATCYDQGVY 136  
QY 59 RKGASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRRRCYQVVG 118  
DB 137 RGTWSTSSGAGCINWNSNLLTRTYNGRRSDAITLGLGNHNYCRNPNNSKPCYVKA 196  
QY 119 LKPLVQECMVHDCPA 132  
DB 197 SKFILEFCSPVCS 210  
RESULT 9  
A35005  
N;Plasminogen activator (EC 3.4.21.73) precursor - chicken  
C;Alternate names: uPA  
C;Species: Gallus gallus (chicken)  
C;Date: 20-Jul-1990 #sequence\_revision 20-Jul-1990 #text\_change 16-Jul-1999  
C;Accession: A35005  
R;Leslie, N.D.; Kessler, C.A.; Bell, S.M.; Degen, J.L.  
J. Biol. Chem. 265, 1339-1344, 1990  
A;Title: The chicken urokinase-type plasminogen activator gene.  
A;Reference number: A35005; MUID:90110185; PMID:2295632  
A;Accession: A35005  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-434 <LES>  
A;Cross-references: GB:J05187; NID:g212858; PIDN:AAA49131.1; PID:g212859  
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try  
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;22-171/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>  
F;40-71/Domain: EGF homology <EGF>  
F;79-158/Domain: kringle homology <KRG>  
F;173-428/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>  
F;173-416/Domain: trypsin homology <TRY>  
F;162-296,202-218,210-285,310-379,342-358,369-397/Disulfide bonds: #status predicted  
F;217,272,373/Active site: His, Asp, Ser #status predicted  
Query Match 42.2%; Score 334.5; DB 1; Length 434;  
Best Local Similarity 54.2%; Pred. No. 7.7e-23;  
Matches 58; Conservative 14; Mismatches 30; Indels 5; Gaps 2;  
QY 11 CDCLNGTGVSNKYFNIHWCNPKKFGGQHCEIDKSKTCYEGNGHFYRGKASTDTMGRP 70

DB 40 CQCLNGTGVSNKYFNIHWCNPKKFGGQHCEIDKSKTCYEGNGHFYRGKASTDTMGRP 95  
QY 71 CLPWNASTVLO-QTYHAHRSALQGLGKHNYCRNPNRRRRCYQVVG 116  
DB 96 CLYNDHPSVIRWGDTHADKALQGLGKHNYCRNPNRRRRCYQVVG 142  
RESULT 10  
I38098  
N;Plasminogen activator precursor, inactive endothelial splice form - human  
C;Alternate names: tissue plasminogen activator  
C;Species: Homo sapiens (man)  
C;Date: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 22-Jun-1999  
C;Accession: I38098; S01678  
R;Siebert, P.D.; Fong, K.  
Nucleic Acids Res. 18, 1086, 1990  
A;Title: Variant tissue type plasminogen activator (PLAT) cDNA obtained from h  
A;Reference number: I38098; MUID:90192128; PMID:1969145  
A;Accession: I38098  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-291 <SIE>  
A;Cross-references: EMBL:X13097; NID:g35282; PIDN:CAA31489.1; PID:g35283  
C;Comment: For the main splice form, see PIR:UKHUT. This form probably does not  
C;Genetics:  
A;Gene: GDB:PLAT  
A;Cross-references: GDB:119496; OMIM:173370  
A;Map position: sp12-ep12  
A;Introns: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2  
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I  
C;Keywords: alternative splicing; fibrinolysis; glycoprotein; kringle  
F;1-23/Domain: signal sequence #status predicted <SIG>  
F;24-32/Domain: propeptide #status predicted <PRO>  
F;33-281/Product: t-plasminogen activator, inactive endothelial splice form #s  
F;41-78/Domain: fibronectin type I repeat homology <EFG>  
F;86-119/Domain: EGF homology <EGF>  
F;127-208/Domain: kringle homology <KRI>  
F;215-291/Domain: kringle homology #status atypical <KR2>  
F;41-71,69-78,86-97,91-108,110-119,127-208,148-190,179-203/Disulfide bonds: #s  
Query Match 41.4%; Score 328.5; DB 2; Length 291;  
Best Local Similarity 46.3%; Pred. No. 1.9e-22;  
Matches 63; Conservative 13; Mismatches 55; Indels 5; Gaps 3  
QY 3 ELHQVP-SNCD---CLNGTGVSNKYFNIHWCNPKKFGGQHCEIDKSKTCYEGNGHFY 58  
DB 77 QCHSVPVKSCSEPRCFNGTCCQALYFSDP-VCCQPEGFAGKCEIDTRATCYEDQGISY 135  
QY 59 RKGASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRRRCYQVVG 118  
DB 136 RGTWSTAESGAECTNWNSSALQNAYSGRRPDAIRLGLGNHNYCRAPDRDSKPCYVKA 195  
QY 119 LKPLVQECMVHDCADG 134  
DB 196 GKYSSEFCSTPACSEG 211  
RESULT 11  
J050599  
N;Plasminogen activator (EC 3.4.21.68) beta precursor - common vampire bat  
C;Alternate names: tissue plasminogen activator  
C;Species: Desmodus rotundus (common vampire bat)  
C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999  
C;Accession: J050599  
R;Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alago  
Gene 105, 229-237, 1991  
A;Title: The plasminogen activator family from the salivary gland of the vampi  
A;Reference number: J050597; MUID:92039036; PMID:1937019  
A;Accession: J050599  
A;Molecule type: mRNA  
A;Residues: 1-431 <KRA>  
A;Cross-references: GB:M63989; NID:g166076; PIDN:AAA31594.1; PID:g166077  
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I

C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;22-36/Domain: propeptide #status predicted <PRO>  
F;37-431/Product: plasminogen activator beta #status predicted <PLA>  
F;41-74/Domain: EGF homology <EGF>  
F;82-163/Domain: kringle homology <KR>  
F;180-425/Domain: trypsin homology <TRY>  
F;41-52,46-63,65-74,82-163,103-145,134-158,168-299,211-227,219-288,313-388/Disulfide bond  
F;139,352/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;179-180/Cleavage site: His-Ser (plasmin) #status predicted  
F;226,275,382/Active site: His, Asp, Ser #status predicted  
F;345-361,378-406/Disulfide bonds: #status predicted

Query Match 41.4%; Score 328.5; DB 2; Length 431;  
Best Local Similarity 47.6%; Pred. No. 2.7e-22;  
Matches 59; Conservative 16; Mismatches 48; Indels 1; Gaps 1;  
QY 9 SNCDCLNGTGVSNKYFVNIHWCNPKPGGQHCIDSKTCYEGNGHFYRGKASTDRMG 68  
Db 42 SEILRCFNGTGWCAASFSDF-VCCPKGYTGQCEVDHATCYKQGVYRGTSWTSBSG 100  
QY 69 RPLCPMNSATVLCQTYHAHRSDALQGLGKHNYCRPNRRPWCYVQVGLKPLVQECMV 128  
Db 101 AQICINNNLLTRTYNGRRSDAITLGLGNHNYCRPNPNNSKPCVVIKASKFILEFCV 160  
QY 129 HDCA 132  
Db 161 PVCS 164

## RESULT 12

UKHTT

N;Alternate names: t-PA; Tissue plasminogen activator  
C;Species: Homo sapiens (man)  
C;Date: 14-Nov-1983 #sequence, revision 14-Nov-1983 #ext5 change 08-Dec-2000  
C;Accession: A94004; A23529; J0562; A93293; S02125; A91343; A93951; A91322; A54645; 160

R;NY, T.; Elgh, F.; Lund, B.  
Proc. Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984  
A;Title: The structure of the human tissue-type plasminogen activator gene: correlation  
A;Reference number: A94004; MUID:84258137; PMID:5089198  
A;Accession: A94004  
A;Molecule type: DNA  
A;Residues: 1-562 <NVT>  
A;Cross-references: GB:L00141

A;Note: The codon given for residue 93 (ACC) is inconsistent with the authors' translation  
R;Friezner Degen, S.J.; Rajput, B.; Reich, E.  
J. Biol. Chem. 261, 6972-6985, 1986  
A;Title: The human tissue plasminogen activator gene.  
A;Reference number: A23529; MUID:86196143; PMID:3009482  
A;Accession: A23529  
A;Molecule type: DNA  
A;Residues: 1-562 <DEG>  
A;Cross-references: GB:X03021; NID:g339817; PIDN:AAA98809.1; PID:g339818

R;Itagaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuoka, S.; Higashio, K.  
Agric. Biol. Chem. 55, 1225-1232, 1991  
A;Title: Purification and characterization of tissue plasminogen activator secreted by H  
A;Reference number: J0562; MUID:91291340; PMID:11368681  
A;Accession: J0562  
A;Molecule type: mRNA  
A;Residues: 31-562 <ITA>  
A;Cross-references: DDBJ:D01096; NID:g220128; PIDN:BA000881.1; PID:g441174

A;Experimental source: embryonic lung fibroblast IMR-90 cells  
A;Note: part of this sequence, including the amino end of the mature protein, was confir  
R;Pennica, D.; Holmes, W.E.; Kohr, W.J.; Harkins, R.N.; Vekari, G.A.; Ward, C.A.; Bennett  
Nature 301, 214-221, 1983  
A;Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Esche  
A;Reference number: A93293; MUID:83115262; PMID:6337343  
A;Accession: A93293  
A;Molecule type: mRNA  
A;Residues: 1-562 <PEN>  
A;Cross-references: GB:L00141  
A;Experimental source: melanoma cells

R;Sasaki, H.; Saito, Y.; Hayashi, M.; Otsuka, K.; Niwa, M.  
Nucleic Acids Res. 16, 5695, 1988  
A;Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA from human  
A;Reference number: S02125; MUID:88262579; PMID:3133640  
A;Accession: S02125  
A;Status: translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-562 <SAS>  
A;Cross-references: EMBL:X07393; NID:g37243; PIDN:CAA30302.1; PID:g37244  
A;Experimental source: fetal lung cells  
R;Kagilani, H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Saito, A.; Bando, H.; Oka, K.;  
FEBS Lett. 189, 145-149, 1985  
A;Title: Expression in Escherichia coli of finger-domain lacking tissue-type plasminogen  
A;Reference number: A91343; MUID:85285620; PMID:3896853  
A;Accession: A91343  
A;Molecule type: mRNA  
A;Residues: 1-38, 'G', 86-433, 'E', 435-562 <KAG>  
A;Experimental source: Detroit 562 cells; ATCC 138  
R;Edlund, T.; Ny, T.; Ranby, M.; Heden, L.O.; Palm, G.; Holmgren, E.; Josephson, S.  
Proc. Natl. Acad. Sci. U.S.A. 80, 349-352, 1983  
A;Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen act  
A;Reference number: A93951; MUID:83169656; PMID:6572897  
A;Accession: A93951  
A;Molecule type: mRNA  
A;Residues: 251-358 <EDL>  
A;Experimental source: melanoma cells  
R;Pohl, G.; Kallstrom, M.; Bergsdorf, N.; Wallen, P.; Jornvall, H.  
Biochemistry 23, 3701-3707, 1984  
A;Title: Tissue plasminogen activator: peptide analyses confirm an indirectly derived  
differences.  
A;Reference number: A90488; MUID:85000468; PMID:6433976  
A;Contents: annotation; melanoma cells, partial sequence of residues 36-562, a live an  
R;Pohl, G.; Kaplan, L.; Einarsson, M.; Wallen, P.; Jornvall, H.  
FEBS Lett. 168, 29-32, 1984  
A;Title: Differences between uterine and melanoma forms of tissue plasminogen  
A;Reference number: A91322; MUID:84158956; PMID:6538514  
A;Accession: A91322  
A;Molecule type: protein  
A;Residues: 33-45, 7311-320 <POH>  
A;Experimental source: uterus  
A;Note: in the uterus, cleavage of the activation peptide may also occur after  
R;van Zonneveld, A.J.; Veerman, H.; Pannekoek, H.  
J. Biol. Chem. 261, 14214-14218, 1986  
A;Reference number: A37567; MUID:87033611; PMID:3021732  
A;Contents: annotation; fibrin binding site  
R;Verheijen, J.H.; Caspers, M.P.M.; Chang, G.T.G.; de Munk, G.A.W.; Pouwels, P.  
EMBO J. 5, 3525-3530, 1986  
A;Title: Involvement of finger domain and kringle 2 domain of tissue-type plas  
A;Reference number: A37568; MUID:87161761; PMID:3030730  
A;Contents: annotation; fibrin binding site  
R;Dodd, I.; Nunn, B.; Robinson, J.H.  
Thromb. Haemost. 59, 523-528, 1988  
A;Title: Isolation, identification and pharmacokinetic properties of human tis  
A;Reference number: A60902; MUID:89044681; PMID:3142086  
A;Contents: annotation; novel forms of expressed recombinant t-PA  
R;Harris, T.J.R.; Patel, T.; Marston, F.A.O.; Little, S.; Emtage, J.S.; Opdena  
Mol. Biol. Med. 3, 279-292, 1986  
A;Title: Cloning of cDNA coding for human tissue-type plasminogen activator an  
A;Reference number: A54645; MUID:86284200; PMID:3090401  
A;Accession: A54645  
A;Molecule type: mRNA  
A;Residues: 1-562 <HAR>  
A;Cross-references: GB:M15518; NID:g190031; PIDN:AAA60111.1; PID:g190032  
A;Note: parts of this sequence were confirmed by peptide sequencing  
R;Reddy, V.B.; Garrazone, A.J.; Sasaki, H.; Wei, C.  
DNA 6, 461-472, 1987  
A;Title: Expression of human uterine tissue-type plasminogen activator in mous  
A;Reference number: I60110; MUID:89054470; PMID:2824147  
A;Accession: I60110  
A;Status: translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-562 <RES>  
A;Cross-references: GB:M18182; NID:g340176; PIDN:AAA36800.1; PID:g340177



R;Fisher, R.; Waller, E.K.; Grossi, G.; Thompson, D.; Tizard, R.; Schleuning, W.D.  
J. Biol. Chem. 260, 11223-11230, 1985  
A;Title: Isolation and characterization of the human tissue-type plasminogen activator cDNA  
A;Reference number: I55232; MUID:85289338; PMID:3161893  
A;Accession: I55232  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-36 <RE2>  
A;Cross-references: GB:M1890; NID:9339837; PIDN:AAA61213.1; PID:9339839  
C;Comment: Cleavage by plasmin or trypsin produces two chains held together by a single Arg-Val bond. 1c  
C;Comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-Val bond. 1c  
C;Comment: t-PA binds chain A of fibrin by kringle 2 and the fibronectin type I repeat.  
C;Genetics:  
A;Gene: GDB:PLAT  
A;Cross-references: GDB:119496; OMIM:173370  
A;Map position: 8p22-8p12  
A;Introns: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2; 297/3; 362/2; 408/1; 455/3; 510/3  
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat homology  
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase  
F;1-23/Domain: signal sequence #status predicted <SIG>  
F;24-32/Domain: propeptide #status predicted <PRO>  
F;33-562/Product: t-plasminogen activator #status experimental <MAT>  
F;33-310/Product: t-plasminogen activator chain A #status experimental <ACH>  
F;41-78/Domain: fibronectin type I repeat homology <IF1>  
F;96-119/Domain: EGF homology <EGF>  
F;127-208/Domain: kringle homology <KR1>  
F;215-296/Domain: kringle homology <KR2>  
F;311-556/Domain: trypsin homology <TRY>  
F;41-71,69,78,86-97,91-108,119,127-208,148-190,179-203,215-296,236-278,267-291,299-455,483/Binding site: carboxylate (Asn) (covalent) #status experimental  
F;219/Binding site: carboxylate (Asn) (covalent) (partial) #status experimental  
F;310-311/Cleavage site: Arg-Ile (plasmin, trypsin) #status experimental  
F;357,406/Active site: His, Asp #status predicted  
F;513/Active site: Ser #status experimental

Query Match 41.4%; Score 328.5; DB 1; Length 562;  
Best Local Similarity 46.3%; Pred. No. 3.3e-22;  
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;  
QY 3 ELHQVP-SNCD---CLNGTCTVSNKYFSNIHWCNCPKFGGHCCEIDKSKTCYEGNGHFY 58  
Db 77 QCHSVPVKSSBPRCFNGGTCQQALYFSDP-VCCQPEGFAGKCEIDTRATCFEGGITY 135  
QY 59 RGKASTDTMGRPCLPNSATVLOOTYHAHRSALQGLGHNYCRPNDRRPPWCYVQVG 118  
Db 136 RGTWSTAEGAECTNNSSALQKPYSGRRPDAILRLGLGNHNYCRNPDREDKPKCVFKA 195  
QY 119 LKPLVQECMVHDCADG 134  
Db 196 GKYSSEFCSTPACSEG 211

RESULT 13  
A35029  
t-plasminogen activator (EC 3.4.21.68) precursor - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C;Accession: A35029; A31597  
F;Feng, P.; Ohlsson, M.; Ny, T.  
J. Biol. Chem. 265, 2022-2027, 1990  
A;Title: The structure of the TATA-less rat tissue-type plasminogen activator gene. Spec  
A;Reference number: A35029; MUID:90130448; PMID:2105315  
A;Accession: A35029  
A;Status: Preliminary  
A;Molecule type: DNA  
A;Residues: 1-559 <FEN>  
A;Cross-references: GB:M31197; NID:9207429; PIDN:AAA42261.1; PID:9207431; GB:J05226  
R;Ny, T.; Leonardsson, G.; Hsueh, A.J.W.  
DNA 7, 671-677, 1988  
A;Title: Cloning and characterization of a cDNA for rat tissue-type plasminogen activator  
A;Reference number: A31597; MUID:89170114; PMID:31148445  
A;Accession: A31597

A;Molecule type: mRNA  
A;Residues: 1-379; 'K' 381-559 <NYT>  
A;Cross-references: GB:M23697; NID:9530159; PIDN:AAA41812.1; PID:9530160  
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat  
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F;1-17/Domain: signal sequence #status predicted <SIG>  
F;18-29/Domain: propeptide #status predicted <PRO>  
F;30-559/Product: t-plasminogen activator #status predicted <MAT>  
F;30-308/Product: t-plasminogen activator chain A #status predicted <ACH>  
F;38-75/Domain: fibronectin type I repeat homology <IF1>  
F;83-116/Domain: EGF homology <EGF>  
F;124-205/Domain: kringle homology <KR1>  
F;213-294/Domain: kringle homology <KR2>  
F;309-559/Product: t-plasminogen activator chain B #status predicted <BCH>  
F;309-553/Domain: trypsin homology <TRY>  
F;38-68,66-75,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-299,297-308,309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted  
F;308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted  
F;335,404,510/Active site: His, Asp, Ser #status predicted  
Query Match 40.5%; Score 321.5; DB 1; Length 559;  
Best Local Similarity 45.6%; Pred. No. 1.4e-21;  
Matches 62; Conservative 14; Mismatches 55; Indels 5; Gaps 2;  
QY 3 ELHQVP-SNCDCLNGGTCVSNKYFSNIHWCNCPKFGGHCCEIDKSKTCYEGNGHFY 58  
Db 74 QCHSVPVKSSBPRCFNGGTCQQALYFSDP-VCCQPEGFAGKCEIDTRATCFEGGITY 132  
QY 59 RGKASTDTMGRPCLPNSATVLOOTYHAHRSALQGLGHNYCRPNDRRPPWCYVQVG 118  
Db 133 RGTWSTAEGAECTNNSSALQKPYSGRRPDAILRLGLGNHNYCRNPDREDKPKCVFKA 192  
QY 119 LKPLVQECMVHDCADG 134  
Db 193 GKYSSEFCSTPACPKG 208

RESULT 14  
A29941  
t-plasminogen activator (EC 3.4.21.68) precursor - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C;Accession: A29941; S48205; S48207; S48206  
R;Rickles, R.J.; Darrow, A.L.; Strickland, S.  
J. Biol. Chem. 263, 1563-1569, 1988  
A;Title: Molecular cloning of complementary DNA to mouse tissue plasminogen ac  
A;Reference number: A29941; MUID:88087303; PMID:2826484  
A;Accession: A29941  
A;Molecule type: mRNA  
A;Residues: 1-559 <RIC>  
A;Cross-references: GB:J03520; NID:9202109; PIDN:AAA40470.1; PID:9202110  
R;Lijnen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.  
Eur. J. Biochem. 224, 863-871, 1994  
A;Title: Characterization of the murine plasma fibrinolytic system.  
A;Reference number: S48202; MUID:95010076; PMID:7523120  
A;Accession: S48205  
A;Molecule type: protein  
A;Residues: 33-37, 'X', 39-40 <LIJ>  
A;Accession: S48207  
A;Molecule type: protein  
A;Residues: 309-316 <LI2>  
A;Accession: S48206  
A;Molecule type: protein  
A;Residues: 33-37, 'X', 39-40 <LIW>  
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat  
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F;1-17/Domain: signal sequence #status predicted <SIG>  
F;18-29/Domain: propeptide #status predicted <PRO>  
F;30-559/Product: t-plasminogen activator #status predicted <MAT>  
F;30-308/Product: t-plasminogen activator chain A #status predicted <ACH>  
F;38-75/Domain: fibronectin type I repeat homology <IF1>  
F;83-116/Domain: EGF homology <EGF>  
F;124-205/Domain: kringle homology <KR1>





F:447,497,598/Active site: His, Asp, Ser #status predicted

Query Match 32.8%; Score 260.5; DB 1; Length 655;  
Best Local Similarity 43.7%; Pred. No. 5.2e-16;  
Matches 52; Conservative 8; Mismatches 42; Indels 17; Gaps 2;  
QY 5 HQVPSNCDLNGGTCVSNKYFSNIHW-----CNCPKKFGGQHCIDKSKTCYEGNGH 56  
Db 242 HTACLSPPCLNGGTC-----HLIVATGTTVCACPPGPGAGRLCNIEPDERCFLNGT 292  
QY 57 FYRGKASDTWGRPCLPWNSATVLQOQTHAHSRDLQGLGKHNCRPNRRPWCYV 115  
Db 293 GYRGVASTASGLSCLAWSLLIQELHVDVSGAALLGLGPHAYCRPNDRPWCYV 351

## RESULT 18

JC4795  
plasma hyaluronan-binding protein precursor - human  
N/Alternate names: hepatocyte growth factor activator-like protein; PHBP  
N/Contains: serine proteinase (EC 3.4.21.-)  
C/Species: Homo sapiens (man)  
C/Date: 15-Oct-1995 #sequence\_revision 16-Aug-1996 #text\_change 19-Jul-2002  
C/Accession: JC4795  
R/Choi-Mura, N.H.; Tobe, T.; Sumiya, J.; Nakano, Y.; Sano, Y.; Mazda, M.  
J. Biochem. 119, 1157-1165, 1996  
A/Title: Purification and characterization of a novel hyaluronan-binding protein (PHBP) r activator.  
A/Reference number: JC4795; MUID:96425001; PMID:8827452  
A/Accession: JC4795  
A/Molecule type: mRNA  
A/Residues: 1-560 <CHO>  
A/Cross-references: GB:S83182; NID:g1836158; PIDN:AA646909.1; PID:g1836159  
A/Experimental source: plasma  
A/Note: parts of this sequence, including the amino ends of the mature chains, were detected

C/Genetics:  
A/Gene: GDB:HABP2; HABP; PHBP; HGFAL  
A/Cross-references: GDB:4573962  
C/Complex: a disulfide-bonded heterodimer of chains produced from the same precursor; the C-subunit is a plasma hyaluronan-binding protein; EGF homology; kringle homology; trypsin; C/Keywords: chondroitin sulfate proteoglycan; glycoprotein; hyaluronic acid; hydrolase; F:1-23/Domain: signal sequence #status predicted <SIG>  
F:24-313/Product: plasma hyaluronan-binding protein, 50K chain #status predicted <50K>  
F:77-108/Domain: EGF homology <EG1>  
F:115-147/Domain: EGF homology <EG2>  
F:154-187/Domain: EGF homology <EG3>  
F:194-276/Domain: kringle homology <KRI>  
F:314-550/Domain: trypsin homology <TRY>  
F:314-516/Product: plasma hyaluronan-binding protein, catalytic chain #status predicted  
F:54,207/Binding site: carboxydrate (Asn) (covalent) #status predicted  
F:77-88,82-97,98-105,115-125,120-136,138-147,154-165,169-176,178-187,194-276,215-257,246  
F:362,405,509/Active site: His, Asp, Ser #status predicted

Query Match 32.5%; Score 257.5; DB 1; Length 560;  
Best Local Similarity 42.1%; Pred. No. 8.5e-16;  
Matches 51; Conservative 18; Mismatches 49; Indels 3; Gaps 3;  
QY 13 CLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIDKSKTCYEGNGHFRYRGKASDTWGRPCL 72  
Db 159 CQNGATCSRHKRSKF-TACPDQPKFGFCEIG-SDDCVGSDGYRGKMRTRVQHAHL 216  
QY 73 PWSATVLQOQTHAHSRDLQGLGKHNCRPNRRPWCYVQGLKPLVQE-CMWHDC 131  
Db 217 YWSHLLLOENYNMFMEDAETHGHEHNFRCNPDADEKPCWCFIKVNDKVKWEYCDVSAC 276  
QY 132 A 132  
Db 277 S 277

## RESULT 19

S28941  
coagulation factor XIIa (EC 3.4.21.38) - guinea pig (fragment)  
N/Alternate names: Hageman factor

C/Species: Cavia porcellus (guinea pig)  
C/Date: 25-Feb-1994 #sequence\_revision 03-Aug-1995 #text\_change 21-Jan-2000  
C/Accession: S28941  
R/Semba, U.; Yamamoto, T.; Kunisada, T.; Shibuya, Y.; Tanase, S.; Kambata, T.; Kabe, Biochim. Biophys. Acta 1159, 113-121, 1992  
A/Title: Primary structure of guinea-pig Hageman factor: sequence around the catalytic site  
A/Reference number: S28941; MUID:93003367; PMID:1390917  
A/Accession: S28941  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-603 <SEM>  
A/Cross-references: EMBL:X68615; NID:949578; PIDN:CAA48600.1; PID:g49579  
C/Suprafamily: coagulation factor XII; EGF homology; fibronectin type I repeat  
C/Keywords: hydrolyase; serine proteinase  
F:46-87/Domain: fibronectin type II repeat homology <IF2>  
F:134-169/Domain: fibronectin type I repeat homology <FB1>  
F:177-208/Domain: EGF homology <EGF>  
F:216-294/Domain: kringle homology <KRG>  
F:359-597/Domain: trypsin homology <TRY>

Query Match 32.4%; Score 257; DB 2; Length 603;  
Best Local Similarity 40.8%; Pred. No. 1e-15;  
Matches 49; Conservative 17; Mismatches 46; Indels 8; Gaps 3

QY 13 CLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIDKSKTCYEGNGHFRYRGKASDTWGRPCL 72  
Db 182 CLNGGRCLE---VEGHLLCDCPMGYTGPFCDLDTTASCYEGRGVSRYGVARITVSGAKCQ 238  
QY 73 PWSATVLQOQTHAHSRDLQGLGKHNCRPNRRPWCYVQGLKPLVQECMWHDC 131  
Db 239 RWAS----BATVNTAEQALRGHHTFCRPNDRPWCYVFMVNGNLSMEYCDLAQC 294

## RESULT 20

KFHUI2  
coagulation factor XIIa (EC 3.4.21.38) precursor [validated] - human  
N/Alternate names: Hageman factor (activated)  
C/Species: Homo sapiens (man)  
C/Date: 27-Nov-1985 #sequence\_revision 30-Jun-1991 #text\_change 08-Dec-2000  
C/Accession: A29411; A26814; A00930; A25191; A22248; A21037  
R/Cool, D.E.; MacGillivray, R.T.A.  
J. Biol. Chem. 262, 13662-13673, 1987  
A/Title: Characterization of the human blood coagulation factor XII gene. Intronic  
A/Reference number: A29411; MUID:88007593; PMID:2888762  
A/Accession: A29411  
A/Molecule type: DNA  
A/Residues: 1-615 <COO>  
A/Cross-references: GB:M17466; GB:J02807; NID:g180355; PIDN:AA859450.1; PID:g180357  
R/Tripodi, M.; Citarella, F.; Guida, S.; Galeffi, P.; Fantoni, A.; Cortese, R. Nucleic Acids Res. 14, 3146, 1986  
A/Title: cDNA sequence coding for human coagulation factor XII (Hageman).  
A/Reference number: A26814; MUID:86176794; PMID:3754331  
A/Accession: A26814  
A/Molecule type: mRNA  
A/Residues: 4-615 <TRI>  
A/Cross-references: GB:M31315; NID:g182291; PIDN:AAA70225.1; PID:g182292  
R/Cool, D.E.; Edgell, C.J.S.; Louie, G.V.; Zoller, M.J.; Brayer, G.D.; MacGillivray, R. J. Biol. Chem. 260, 13666-13676, 1985  
A/Title: Characterization of human blood coagulation factor XII cDNA. Prediction  
A/Reference number: A00930; MUID:86033830; PMID:3877053  
A/Accession: A00930  
A/Molecule type: mRNA  
A/Residues: 14-332, S', 334-615 <CO2>  
A/Cross-references: GB:M11723; NID:g180358; PIDN:AAA51986.1; PID:g180359  
R/Que, B.G.; Davie, E.W. Biochemistry 25, 1525-1528, 1986  
A/Title: Characterization of a cDNA coding for human factor XII (Hageman factor).  
A/Reference number: A25191; MUID:86216049; PMID:3011063  
A/Accession: A25191  
A/Molecule type: mRNA  
A/Residues: 146-378, G', 380-615 <QUE>  
A/Cross-references: GB:M13147; NID:g180360; PIDN:AAA70224.1; PID:g180361  
R/McMullen, B.A.; Fujikawa, K.

J. Biol. Chem. 260, 5328-5341, 1985  
A;Title: Amino acid sequence of the heavy chain of human alpha-factor XIIa (activated Ha  
A;Reference number: A22248; MUID:85182674; PMID:3886654  
A;Accession: A22248  
A;Molecule type: protein  
A;Residues: 20-379 <MCW>  
R;Fujikawa, K.; McMullen, B.A.  
J. Biol. Chem. 258, 10924-10933, 1983  
A;Title: Amino acid sequence of human beta-factor XIIa.  
A;Reference number: A21037; MUID:83291041; PMID:6604055  
A;Accession: A21037  
A;Molecule type: protein  
A;Residues: 354-362;373-615 <FU>  
R;Harris, R.J.; Ling, V.T.; Spellman, M.W.  
J. Biol. Chem. 267, 5102-5107, 1992  
A;Title: O-linked fucose is present in the first epidermal growth factor domain of factor  
A;Reference number: A44606; MUID:92184750; PMID:1544894  
C;Genetics:  
A;Gene: GDB:F12  
A;Cross-references: GDB:119892; OMIM:234000  
A;Map position: 5q34-qter  
A;Introns: 19/3; 39/1; 72/2; 96/1; 133/1; 267/2; 340/1; 417/2; 463/1; 511/  
C;Complex: factor XII, prekallikrein, and HMW kininogen form a complex bound to anionic  
C;Function:  
A;Description: factor XIIa catalyzes the proteolytic activation of plasminogen, plasma p  
ixrein  
A;Pathway: blood coagulation; fibrinolysis  
C;Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;  
C;Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; kringle; plasma;  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-372,373-615/Product: coagulation factor XIIa, alpha form #status experimental <A12>  
F;47-88/Domain: fibronectin type II repeat homology <FB2>  
F;98-130/Domain: EGF homology <EG1>  
F;135-170/Domain: fibronectin type I repeat homology <FI1>  
F;178-209/Domain: EGF homology <EG2>  
F;217-295/Domain: kringle homology <KRG>  
F;298-356/Region: proline-rich  
F;354-362,373-615/Product: coagulation factor XIIa, beta form #status experimental <B12>  
F;373-609/Domain: trypsin homology <TRY>  
F;98-110,104-119,121-130,135-163,161-170,178-189,183-198,200-217,295,238-277,266-290  
F;109/Binding site: carbohydrate (Thr) (covalent) #status experimental  
F;249,433/Binding site: carbohydrate (Thr) (covalent) #status experimental  
F;298,305,329,337/Binding site: carbohydrate (Thr) (covalent) #status predicted  
F;308/Binding site: carbohydrate (Ser) (covalent) #status predicted  
F;412,461,563/Active site: His, Asp, Ser #status predicted  
Query Match 29.4%; Score 233; DB 1; Length 615;  
Best Local Similarity 42.3%; Pred. No. 1.5e-13;  
Matches 44; Conservative 12; Mismatches 40; Indels 8; Gaps 3;  
CY 13 CLNGGTCVSNKYFSNIHWNCNPKKFGQCHCEIDKSKTCYEGNGHYRKGKASTDTWGRPCL 72  
DB 183 CLHGGRCLE--VEGHRLCPCPVGYTGPFCDVDTKASCYDGRGLSYRGLARTTLGAPCQ 239  
CY 73 PWNASATVLOQTY-HAHRSDALQLGLGKHNYCRNPNRRPWCYV 115  
DB 240 PWAS-----EATYRNVTAQRNWLGLGHAFCRNPDNDRPWCYV 279  
RESULT 21  
t-plasminogen activator (EC 3.4.21.68) gamma precursor - common vampire bat  
N;Alternate names: tissue plasminogen activator  
C;Species: Desmodus rotundus (common vampire bat)  
C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999  
C;Accession: J05600  
R;Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Don  
Gene 105, 229-237, 1991  
A;Title: the plasminogen activator family from the salivary gland of the vampire bat Des  
A;Reference number: J05600; MUID:92039036; PMID:1937019  
A;Accession: J05600  
A;Molecule type: mRNA

A;Residues: 1-394 <KRA>  
A;Cross-references: GB:M63990; MUID:9166078; PIDN:AAA31595.1; PID:9166079  
A;Note: the authors translated the codon ATC for residue 75 as Thr  
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I  
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;22-36/Domain: propeptide #status predicted <PRO>  
F;45-126/Domain: plasminogen activator gamma #status predicted <PLA>  
F;143-388/Domain: trypsin homology <TRY>  
F;145-126,66-108,97-121,131-262,174-190,182-251,276-351,308-324,341-369/Disulfide  
F;142-143/Cleavage site: His-Ser (plasmin) #status predicted  
F;189,238,345/Active site: His, Asp, Ser #status predicted  
F;315/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 28.1%; Score 223; DB 2; Length 394;  
Best Local Similarity 44.3%; Pred. No. 8.2e-13;  
Matches 39; Conservative 13; Mismatches 36; Indels 0; Gaps 0  
CY 45 DKSKTCYEGNGHYRKGKASTDTWGRPCLPWNASATVLOQTYHAHRSDALQLGLGKHNYCRN 104  
DB 40 DPHATCYKQGVYRTWTSTSSGACINWNNLLIRTYNGRMPENAVKGLGKHNYCRN 99  
CY 105 PNRRRPWCYVQVGLKPLVQECMWHDC 132  
DB 100 PDGASKPCWCVIKARKFTSESCSPVCS 127  
RESULT 22  
S45281  
coagulation factor XIIa (EC 3.4.21.38) precursor - bovine (fragment)  
N;Alternate names: Hageman factor (activated)  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 10-Apr-1995 #sequence\_revision 22-Apr-1995 #text\_change 21-Jan-2000  
C;Accession: S45281; A61329  
R;Shibuya, Y.; Samba, U.; Okabe, H.; Kanbara, T.; Yamamoto, T.  
Biochim. Biophys. Acta 1208, 63-70, 1994  
A;Title: Primary structure of bovine Hageman factor (blood coagulation factor  
A;Reference number: S45281; MUID:94242782; PMID:8186251  
A;Accession: S45281  
A;Molecule type: mRNA  
A;Residues: 1-593 <SHI>  
A;Cross-references: GB:S70164  
A;Note: the authors translated the codon GAG for residue 23 as Val, GAG for re  
is, and ATC for residue 505 as Leu  
R;Fujikawa, K.; Walsh, K.A.; Davie, E.W.  
Biochemistry 16, 2270-2278, 1977  
A;Title: Isolation and characterization of bovine factor XII (Hageman factor).  
A;Reference number: A61329; MUID:77182112; PMID:861210  
A;Accession: A61329  
A;Molecule type: protein  
A;Residues: 10-16, 'X', 18-19,525-550 <FUJ>  
C;Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat  
C;Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; monomer;  
F;37-78/Domain: fibronectin type II repeat homology <IF2>  
F;88-120/Domain: EGF homology <EGF>  
F;125-160/Domain: fibronectin type I repeat homology <FI1>  
F;207-287/Domain: kringle homology <KRG>  
F;350-587/Domain: trypsin homology <TRY>  
F;541/Active site: Ser #status predicted  
Query Match 27.4%; Score 217; DB 2; Length 593;  
Best Local Similarity 35.7%; Pred. No. 4e-12;  
Matches 46; Conservative 15; Mismatches 58; Indels 10; Gaps 4  
CY 6 QVPSNCDLNGTGVSNKYFSNIHWNCNPKKFGQCHCEIDKSKTCYB--GNGHFYRKGAS 63  
DB 166 QVCRTNPLNGSDSLQAE---GHRLCRCAPSFAGRLCDVLDKASCYDGRGLSYRGMAG 222  
CY 64 TDTWGRCLPWNASATVLOQTY-HAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLKPL 122  
DB 223 TLLGAPCQSWAS-----EATYRNVTAQRNWLGLGHAFCRNPDNDRPWCYV 276



F:2650-2727/Domain: kringle homology <KR24>  
F:2764-2841/Domain: kringle homology <KR25>  
F:2878-2955/Domain: kringle homology <KR26>  
F:2992-3069/Domain: kringle homology <KR27>  
F:3106-3183/Domain: kringle homology <KR28>  
F:3220-3297/Domain: kringle homology <KR29>  
F:3334-3411/Domain: kringle homology <KR30>  
F:3448-3525/Domain: kringle homology <KR31>  
F:3562-3639/Domain: kringle homology <KR32>  
F:3676-3753/Domain: kringle homology <KR33>  
F:3782-3859/Domain: kringle homology <KR34>  
F:3896-3973/Domain: kringle homology <KR35>  
F:4010-4087/Domain: kringle homology <KR36>  
F:4124-4201/Domain: kringle homology <KR37>  
F:4228-4307/Domain: kringle homology <KR38>  
F:4328-4541/Domain: trypsin homology <TRY>  
  
Query Match 20.5%; Score 162.5; DB 1; Length 4548;  
Best Local Similarity 33.1%; Pred. No. 1.8e-06; Indels 25; Gaps 7;  
Matches 41; Conservative 12; Mismatches 46;  
  
QY 27 NIHW--CN-----KFGGQHCHIDSKT---CYEGNGHYRGKASTDTMGR 69  
DB 3742 NVREYCNLTQCPVTSESSVLTATSTAVSEQAPTEQSPVTQDCYHGQSGSYRGSFTTIVTGR 3801  
  
QY 70 PCLPWSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLKPLV--QECM 127  
DB 3802 TCQSSMT---PHWQRTTEYPNGGLTRNYCRNPDAEIRPWCYT---MDPSVRWEYCN 3855  
  
QY 128 VHDC 131  
DB 3856 LTQC 3859  
  
RESULT 25  
A32869  
apolipoprotein(a) (EC 3.4.21.-) - rhesus macaque (fragment)  
C:Species: Macaca mulatta (rhesus macaque)  
C:Date: 22-Nov-1989 #sequence\_revision 22-Nov-1989 #text\_change 22-Jun-1999  
C:Accession: A32869; A30848  
J:Tomlinson, J.E.; McLean, J.W.; Lawn, R.M.  
J. Biol. Chem. 264, 5957-5965, 1989  
A:Title: Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of synthesis.  
A:Reference number: A32869; MUID:89174660; PMID:2925643  
A:Accession: A32869  
A:Molecule type: mRNA  
A:Residues: 1-1420 <TOM>  
A:Cross-references: GB:J04635; NID:9342072; PIDN:AAA36833.1; PID:9342073  
C:Superfamily: apolipoprotein(a); kringle homology; trypsin homology  
C:Keywords: hydrolase; kringle; lipid binding; lipoprotein; serine proteinase  
F:50-127/Domain: kringle homology <KR1>  
F:164-241/Domain: kringle homology <KR2>  
F:278-355/Domain: kringle homology <KR3>  
F:392-469/Domain: kringle homology <KR4>  
F:506-583/Domain: kringle homology <KR5>  
F:620-697/Domain: kringle homology <KR6>  
F:726-803/Domain: kringle homology <KR7>  
F:840-917/Domain: kringle homology <KR8>  
F:954-1031/Domain: kringle homology <KR9>  
F:1068-1145/Domain: kringle homology <KR10>  
F:1191-1413/Domain: trypsin homology <TRY>  
  
Query Match 20.3%; Score 161; DB 2; Length 1420;  
Best Local Similarity 40.9%; Pred. No. 9.4e-07; Indels 12; Gaps 4;  
Matches 36; Conservative 9; Mismatches 31;  
  
QY 50 CYEGNGHYRGKASTDTMGSPCLPWSATVLQ--QTYHAHRSALQGLGKHNYCRNPN 107  
DB 1068 CYHNGSYRGSTSTVTGTCQSSMTPHQKRTPEHFNDDLTN-----NYCRNPDA 1122  
  
QY 108 RRPWCYVQVGLKPLVQE--CMVHDCAD 133  
DB 1123 DTGPWCFT---MDPSVRREYCNLTRCSD 1147

## RESULT 26

E61545  
plasmin (EC 3.4.21.7) precursor - goat (fragments)  
N:Alternate names: plasminogen  
C:Species: Capra aegagrus hircus (domestic goat)  
C:Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 12-May-1995  
C:Accession: C61545  
R:Schaller, J.; Rickli, E.E.  
Enzyme 40, 63-69, 1988  
A:Title: Structural aspects of the plasminogen of various species.  
A:Reference number: A61545; MUID:89005015; PMID:3168975  
A:Accession: C61545  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-123 <SCH>  
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor  
C:Keywords: hydrolase; serine proteinase  
F:41-118/Domain: kringle homology <KR4>

Query Match 20.2%; Score 160; DB 2; Length 123;

Best Local Similarity 34.5%; Pred. No. 1.5e-07; Indels 20; Gaps 6;

Matches 39; Conservative 17; Mismatches 35;

QY 35 KRFGGQ-----HCEIDSKTCYEGNGHYRGKASTDTMGSPCLPWSATV--LQQTYH 85

DB 20 KKLGRSVDCAKCE--EAAQCYHNGQSYRGSTSTVTGTRKQSSSMIPHRHOKTPE 78

QY 86 AHESDALQGLGKHNYCRNPNRRPWCYVQVGLKPLV--QECMVHDCAD 133

DB 79 SYPNAGLTM-----NYCRNPDAKSPWCYT---TDPVRWEFCNLKKCS 120

## RESULT 27

E61545  
plasmin (EC 3.4.21.7) precursor - dog (fragments)

N:Alternate names: plasminogen

C:Species: Canis lupus familiaris (dog)

C:Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 12-May-1995

C:Accession: E61545

R:Schaller, J.; Rickli, E.E.

Enzyme 40, 63-69, 1988

A:Title: Structural aspects of the plasminogen of various species.

A:Reference number: A61545; MUID:89005015; PMID:3168975

A:Accession: E61545

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-120 <SCH>

C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor

C:Keywords: hydrolase; serine proteinase

F:37-114/Domain: kringle homology <KR4>

Query Match 20.1%; Score 159; DB 2; Length 120;

Best Local Similarity 38.9%; Pred. No. 1.8e-07;

Matches 37; Conservative 8; Mismatches 32; Indels 18; Gaps 5

QY 46 KSKTCYEGNGHYRGKASTDTMGSPCLPWSATVLQQTYHAHRSDAL----QLGLGKH 100

DB 33 KVQCYHNGQSYRGSTSTITGRKQSSSMT-----PHRHKTPHFPEAGL-TMN 84

QY 101 YCRNPNRRRPWCYVQVGLKPLV--QECMVHDCAD 133

DB 85 YCRNPDAKSPWCYT---TDPVRWEFCNLKCLD 116

## RESULT 28

A45082

neurotrophic receptor rorl precursor - human

N:Contains: protein-tyrosine kinase (EC 2.7.1.112)

C:Species: Homo sapiens (man)

C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 19-May-2000

C:Accession: A45082

R;Maslakowski, P.; Carroll, R.D.  
J. Biol. Chem. 267, 26181-26190, 1992  
A;Title: A novel family of cell surface receptors with tyrosine kinase-like domain.  
A;Reference number: A45082; MUID:93100347; PMID:1334494  
A;Accession: A45082  
A;Molecule type: mRNA  
A;Residues: 1-937 <MAS>  
A;Cross-references: GB:M97675; NID:G337464; PIDN:AAA60275.1; PID:G337465  
A;Note: sequence extracted from NCBI backbone (NCBIP:120916)  
C;Genetics:  
A;Gene: GDB:NTRKRI  
A;Cross-references: GDB:136453  
A;Map position: 6p21-6p21  
C;Superfamily: neurotrophic receptor ror; immunoglobulin homology; kringle homology; pro  
C;Keywords: ATP; glycoprotein; kringle; phosphotransferase; transmembrane protein; tyro  
F;1-23/Domain: signal sequence #status predicted <SIG>  
F;24-937/Product: neurotrophic receptor rorl #status predicted <MAT>  
F;72-133/Domain: immunoglobulin homology <IMW>  
F;733-931/Domain: kringle homology <KKG>  
F;404-425/Domain: transmembrane #status predicted <TML>  
F;471-753/Domain: protein kinase homology <KIN>  
F;47,66,184,315/Binding site: carboxylate (Asn) (covalent) #status predicted  
Query Match 20.0%; Score 158.5; DB 2; Length 937;  
Best Local Similarity 29.7%; Pred. No. 1.1e-06;  
Matches 43; Conservative 14; Mismatches 51; Indels 37; Gaps 7;  
QY 2 NELHQVP-----SNDCLENGTGVSNKYS-----NIHCNC-----PKFGGQH 41  
DB 236 DETSVPRDLRDECEILENLVLCQTEYIFARSNPMLMLKLPNCEDLPQPSPEAAN 295  
QY 42 C-----RIDSKTCYEGNGHYRGKASTDTMGPRCLPWNATVLOQTYHAHRSAL 92  
DB 286 CIRIGIPVADPINKHKYNTGVDYGTGVSQRCQPNWS-----QVPHTHTTAL 350  
QY 93 QLQ--LGKHNVCNPNRR--PCWCY 114  
DB 351 RPPELNGHSGVCRNPNQKEAPWCF 375  
RESULT 29  
B61545  
Plasmin (EC 3.4.21.7) precursor - sheep (fragments)  
N;Alternate names: plasminogen  
C;Contains: miniplasminogen  
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C;Date: 28-Oct-1994 #sequence\_revision 01-Nov-1996 #text\_change 17-Mar-1999  
A;Cross-references: B61545; S28200  
R;Schaller, J.; Rickli, E.E.  
Enzyme 40, 63-69, 1988  
A;Title: Structural aspects of the plasminogen of various species.  
A;Reference number: A61545; MUID:89005015; PMID:3168975  
A;Accession: B61545  
A;Molecule type: protein  
A;Residues: 1-37;38-117 <SCH>  
R;Schaller, J.; Straub, C.; Kaempfer, U.; Rickli, E.E.  
Protein Seq. Data Anal. 5, 21-25, 1992  
A;Title: Complete amino acid sequence of ovine miniplasminogen.  
A;Reference number: S28200; MUID:93149995; PMID:1492092  
A;Accession: S28200  
A;Molecule type: protein  
A;Residues: 118-460 <SCZ>  
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology  
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase; z  
F;1-37,38-117,118-460/Product: plasminogen (fragments) #status experimental <PRO>  
F;1-37/Domain: activation peptide (fragment) #status experimental <APT>  
F;38-117,118-230,231-460/Product: plasmin (fragments) #status experimental <MAT>  
F;41-118/Domain: kringle homology <KR4>  
F;118-460/Product: miniplasminogen #status experimental <MIN>  
F;132-211/Domain: kringle homology <KR5>  
F;226-460/Domain: plasmin chain B #status experimental <BCH>  
F;231-453/Domain: trypsin homology <TRY>  
F;272,315,410/Active site: His, Asp, Ser #status predicted

Query Match 19.9%; Score 157.5; DB 2; Length 460;  
Best Local Similarity 38.2%; Pred. No. 7.6e-07;  
Matches 34; Conservative 13; Mismatches 27; Indels 15; Gaps 4  
QY 35 KFGGQ-----HCEIDSKTCYEGNGHYRGKASTDTMGPRCLPWNATV--LQQTYH 85  
DB 20 KKLGRSVEDCAKCE--EERQDCYHNGOYRGTSSTVTGKQCQSWSMIPHRHQKTPE 78  
QY 86 AHRSDALQGLGKHNVCNPNRR--PCWCY 114  
DB 79 SYPNAGLTM-----NYCNPFDADKSPWCY 102  
RESULT 30  
PLBO  
Plasmin (EC 3.4.21.7) precursor - bovine  
N;Alternate names: plasminogen  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 30-Sep-1987 #sequence\_revision 28-Apr-1995 #text\_change 18-Jun-1999  
C;Accession: S45046; A25835; I45961; S03736  
R;Berglund, L.; Andersen, M.D.; Petersen, T.E.  
submitted to the EMBL Data Library, May 1994  
A;Description: Cloning and characterization of the bovine plasminogen cDNA.  
A;Reference number: S45046  
A;Accession: S45046  
A;Molecule type: mRNA  
A;Residues: 1-812 <BER>  
A;Cross-references: EMBL:X79402; NID:G494962; PIDN:CAAS5939.1; PID:G494963  
A;Experimental source: liver  
A;Note: it is uncertain whether Met-1 or Met-8 is the initiator  
R;Schaller, J.; Moser, P.W.; Danneberg-Muller, G.A.K.; Rosselet, S.J.; Kampfer  
Eur. J. Biochem. 149, 267-278, 1985  
A;Title: Complete amino acid sequence of bovine plasminogen. Comparison with h  
A;Reference number: A25835; MUID:85203906; PMID:364532  
A;Accession: A25835  
A;Molecule type: protein  
A;Residues: 27-334,'D',336-515,'H',517-554,'L',556-812 <SCH>  
R;Malinowski, D.P.; Sadler, J.E.; Davie, E.W.  
Biochemistry 23, 4243-4250, 1984  
A;Title: Characterization of a complementary deoxyribonucleic acid coding for  
A;Reference number: I45961; MUID:85023311; PMID:6148361  
A;Accession: I45961  
A;Status: translated from GB/EMBL/DBDJ  
A;Molecule type: mRNA  
A;Residues: 706-743,'R',745-812 <MAL>  
A;Cross-references: GB:K02935; NID:G163551; PIDN:AAA30714.1; PID:G163552  
R;Brunsholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Mann  
Eur. J. Biochem. 114, 465-470, 1981  
A;Title: Comparison of the primary structure of the N-terminal CNBr fragments  
A;Reference number: S03735; MUID:81212097; PMID:7238497  
A;Accession: S03736  
A;Molecule type: protein  
A;Residues: 27-83 <BRU>  
C;Function:  
A;Description: dissolves the fibrin of blood clots; acts as a proteolytic fact  
ns the walls of the graafian follicle; also activates the urokinase-type plas  
C;Pathway: fibrinolysis  
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor  
C;Keywords: duplication; fibrinolysis; glycoprotein; hydrolase; kidney; krinkl  
E;1-26/Domain: signal sequence #status predicted <SIG>  
E;8-103/Domain: plasminogen-related protein precursor homology <PLPH>  
F;27-812/Product: plasminogen #status experimental <PRO>  
F;27-103/Domain: activation peptide #status experimental <APT>  
F;104-583,584-812/Product: plasmin #status experimental <MAT>  
F;104-583/Domain: plasmin chain A #status experimental <ACH>  
F;110-188/Domain: kringle homology <KR1>  
F;192-269/Domain: kringle homology <KR2>  
F;282-359/Domain: kringle homology <KR3>  
F;384-461/Domain: kringle homology <KR4>  
F;485-564/Domain: kringle homology <KR5>  
F;584-812/Domain: plasmin chain B #status experimental <BCH>  
F;584-805/Domain: trypsin homology <TRY>

F:56-80,60-68,110-188,131-171,159-183,192-269,195-323,213-252,241-264,282-359,303-342,333  
Bonds: #status predicted  
F:315/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F:365/Binding site: carbohydrate (Ser) (covalent) #status experimental  
F:624,667,762/Active site: His, Asp, Ser #status predicted

Query Match 19.7%; Score 156; DB 1; Length 812;  
Best Local Similarity 28.3%; Pred. No. 1.7e-06;  
Matches 47; Conservative 12; Mismatches 53; Indels 54; Gaps 10;

QY 5 HQVPSNCDCLN-----GGTCVSNKYFSNIHW--CNCFKKFGGQHCIDKSKT 49  
DB 315 NRTPEFPCKNLEENYCRNPNGEKAPWCYT--SEVRWEYCTIPS-----CBSSPLST 366  
QY 50 -----CVGNGHFYRGKASTDTMGRCPLPWSATVLQOYVHAH--RS 89  
DB 367 ERMDVPPEQTPVQDCHGNGQSYRTSTTTIGRKCQSWSS-----WTFPHRLKTIPE 421

QY 90 DALQLGLGKHNCRNPDRRPPWCYVQVGLKPLV--QECMVHDCAD 133  
DB 422 NYPNAGL-TWNYCRNPDADKSPWCYT--TDPRVRWEYCNLKKCSE 463

RESULT 31  
PUPG  
Plasmin (EC 3.4.21.7) precursor - pig (fragment)  
N/Alternate names: plasminogen  
N/Contains: miniplasminogen  
C/Species: Sus scrofa domestica (domestic pig)  
C/Date: 07-Sep-1990 #sequence revision 01-Nov-1996 #text\_change 18-Jul-1997  
C/Accession: S03733; S03737; A25834  
R/Schaller, J.; Marti, T.; Roessliet, S.J.; Kaempfer, U.; Rickli, E.B.  
Fibrinolysis 1, 91-102, 1987  
A/Title: Amino acid sequence of the heavy chain of porcine plasmin. Comparison of the ca  
A/Reference number: S03733  
A/Accession: S03733  
A/Molecule type: protein  
A/Residues: 1-560 <SCH>  
R/Brunisholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.B.; Lergier, W.; Manneberg, M.;  
Eur. J. Biochem. 114, 465-470, 1981  
A/Title: Comparison of the primary structure of the N-terminal CNBr fragments of human,  
A/Reference number: S03735; MUID:81212097; PMID:7238497  
A/Accession: S03737  
A/Molecule type: protein  
A/Residues: 1-57 <BRU>  
R/Marti, T.; Schaller, J.; Rickli, E.B.  
Eur. J. Biochem. 149, 279-285, 1985  
A/Title: Determination of the complete amino-acid sequence of porcine miniplasminogen.  
A/Reference number: A25834; MUID:85203907; PMID:3846533  
A/Accession: A25834  
A/Molecule type: protein  
A/Residues: 450-790 <MAR>  
C/Function:  
A/Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a va  
ns the walls of the graafian follicle; also activates the urokinase-type plasminogen act  
A/Pathway: fibrinolysis  
C/Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology  
C/Keywords: fibrinolysis; glycoprotein; hydrolase; kidney; kringle; plasma; serine prote  
F:1-790/Product: plasminogen #status predicted <PRO>  
F:1-77/Domain: plasminogen-related protein precursor homology (fragment) <PLPH>  
F:1-77/Domain: activation peptide #status predicted <ACH>  
F:78-560/Product: plasmin chain A #status predicted <ACH>  
F:84-162/Domain: kringle homology <KR1>  
F:166-243/Domain: kringle homology <KR2>  
F:256-333/Domain: kringle homology <KR3>  
F:358-435/Domain: kringle homology <KR4>  
F:450-790/Product: miniplasminogen #status experimental <MIN>  
F:461-540/Domain: kringle homology <KR5>  
F:561-790/Product: plasmin chain B #status experimental <BCH>  
F:561-783/Domain: trypsin homology <TRY>  
F:30-54,34-42,84-162,105-145,133-157,166-243,169-297,187-226,215-238,256-333,277-316,305  
bonds: #status predicted  
F:602,645,740/Active site: His, Asp, Ser #status predicted

Query Match 19.5%; Score 155; DB 1; Length 790;  
Best Local Similarity 27.2%; Pred. No. 2e-06;  
Matches 44; Conservative 16; Mismatches 56; Indels 46; Gaps 9

QY 5 HQVPSNCDCLNGTCVSNKYFSN-----IHW-CNPKKFGGQHCIDK----- 46  
DB 289 NRTPEFPCKN-----LEENYCRNPDEGTAPWCYT--TDSEVRWDYCKIPSCGSSSTTSTHLD 344  
QY 47 -----SKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOYVHAH--SDALQ 93  
DB 345 APVPPPEQTPVQDCHGNGQSYRTSTTTIGRKCQSW-----VMTPHRHEKTPONFPN 399  
QY 94 LGLGKHNCRNPDRRPPWCYVQVGLKPLV--QECMVHDCAD 133  
DB 400 AGL-TWNYCRNPDADKSPWCYT--TDPRVRWEYCNLKKCSE 437

RESULT 32  
A60140  
plasmin (EC 3.4.21.7) precursor - chicken (fragment)  
N/Alternate names: plasminogen  
C/Species: Gallus gallus (chicken)  
C/Date: 22-Jan-1993 #sequence revision 22-Jan-1993 #text\_change 16-Jul-1999  
C/Accession: A60140  
R/Gyenes, M.; Pathy, L.  
Biochim. Biophys. Acta 832, 326-330, 1985  
A/Title: The kringle 4 domain of chicken plasminogen.  
A/Reference number: A60140; MUID:86077796; PMID:4074753  
A/Accession: A60140  
A/Molecule type: protein  
A/Residues: 1-89 <GYE>  
C/Superfamily: plasmin; kringle homology; plasminogen-related protein precursor  
C/Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine pro  
F:6-83/Domain: kringle homology <KR>  
F:6-83,27-66,55-78/Diulfide bonds: #status predicted  
F:39/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 19.4%; Score 154; DB 2; Length 89;  
Best Local Similarity 37.4%; Pred. No. 3.9e-07;  
Matches 34; Conservative 9; Mismatches 36; Indels 12; Gaps 4

QY 47 SKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOYVHAHSDALQGLG--KHNYCRN 104  
DB 3 TEECYQNGVSYRGTSFTITGKCAWNS-----MSPHRHNKTESHFPNADLRQNYCRN 57

QY 105 PDNRPPWCYVQVGLKPLV--QECMVHDCAD 133  
DB 58 PDADRSPWCYT--TDPSVRWEYCNLKRCS 85

RESULT 33  
PLHU  
plasmin (EC 3.4.21.7) precursor [validated] - human  
N/Alternate names: plasminogen precursor [mismomer]  
N/Contains: angiotatin; microplasmin; plasminogen  
C/Species: Homo sapiens (man)  
C/Date: 24-Apr-1984 #sequence revision 02-Dec-1994 #text\_change 15-Sep-2000  
C/Accession: A35229; I52242; A26646; I62738; I84609; S03735; A00929; A04627; A  
R/Petersen, T.E.; Martzen, M.R.; Ichinose, A.; Davie, E.W.  
J. Biol. Chem. 265, 6104-6111, 1990  
A/Title: Characterization of the gene for human plasminogen, a key proenzyme i  
A/Reference number: A35229; MUID:90202879; PMID:2318848  
A/Accession: A35229  
A/Molecule type: DNA  
A/Residues: 1-810 <PBT>  
A/Cross-references: GB:J05286; GB:M34276; NID:G190064; PIDN:AAA60113.1; PID:G3  
A/Experimental source: leukocyte; lung fibroblast  
R/Margaret, N.; Bruno, L.; Pontoglio, M.; Candiani, G.; Meroni, G.; Ottoleng  
Biochem. Biophys. Res. Commun. 173, 1013-1018, 1990  
A/Title: Definition of the transcription initiation site of human plasminogen  
A/Reference number: I52242; MUID:91097523; PMID:2268308  
A/Accession: I52242

A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-16 <MAL1>  
A:Cross-references: GB:M62890; NID:G190092; PIDN:AAA36454.1; PID:G5533613  
R:Forsgren, M.; Raden, B.; Israelsson, M.; Larsson, K.; Heden, L.O.  
FEBS Lett. 213, 254-260, 1987  
A>Title: Molecular cloning and characterization of a full-length cDNA clone for human plasminogen activator  
A:Reference number: A26646; MUID:87162490; PMID:3030813  
A:Accession: A26646  
A:Molecule type: mRNA  
A:Residues: 1-471, 'D', 473-810 <FOR>  
A:Cross-references: GB:X05199; NID:G35530; PIDN:CAA28831.1; PID:G35531  
A:Experimental source: liver  
R:Malinowski, D.P.; Sadler, J.E.; Davie, E.W.  
Biochemistry 23, 4243-4250, 1984  
A>Title: Characterization of a complementary deoxyribonucleic acid coding for human and human plasminogen activator  
A:Reference number: I45961; MUID:85023311; PMID:6148961  
A:Accession: I62738  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 292-471, 'D', 473-810 <MAL2>  
A:Cross-references: GB:K02922; NID:G190112; PIDN:AAA60124.1; PID:G387031  
A:Accession: I84609  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 367-419 <MAL3>  
A:Cross-references: GB:K02921; NID:G190110; PIDN:AAA60123.1; PID:G190111  
R:Brundisholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M.; Eur. J. Biochem. 114, 465-470, 1981  
A>Title: Comparison of the primary structure of the N-terminal CNBr fragments of human, human plasminogen activator and human plasminogen activator  
A:Reference number: S03735; MUID:81212097; PMID:7238497  
A:Accession: S03735  
A:Molecule type: protein  
A:Residues: 20-71, 'E', 73-76 <BRU>  
R:Sottrup-Jensen, L.; Petersen, T.E.; Magnusson, S.  
Submitted to the Atlas, July 1977  
A:Reference number: A00929  
A:Accession: A00929  
A:Molecule type: protein  
A:Residues: 20-71, 'E', 73-85, 87-106, 'D', 108-360, 'E', 362-810 <SOT>  
R:Wiman, B.  
Eur. J. Biochem. 76, 129-137, 1977  
A>Title: Primary structure of the B-chain of human plasmin.  
A:Reference number: A04627; MUID:77225245; PMID:142009  
A:Accession: A04627  
A:Molecule type: protein  
A:Residues: 581-810 <W11>  
R:Wiman, B.; Wallen, P.  
Eur. J. Biochem. 50, 489-494, 1975  
A>Title: Structural relationship between "glutamic acid" and "lysine" forms of human plasminogen activator  
A:Reference number: A04625; MUID:75093329; PMID:122932  
A:Accession: A04625  
A:Molecule type: protein  
A:Residues: 20-50, 'Q', 51-71, 'E', 73-85, 87-100 <W12>  
R:Wiman, B.; Wallen, P.  
Eur. J. Biochem. 58, 539-547, 1975  
A>Title: Amino-acid sequence of the cyanogen-bromide fragment from human plasminogen the fragment from human plasminogen  
A:Reference number: A04626; MUID:76043692; PMID:126863  
A:Accession: A04626  
A:Molecule type: protein  
A:Residues: 483-507, 'E', 509-604 <W13>  
R:Robbins, K.C.; Bernabe, P.; Arzadon, L.; Summaria, L.  
J. Biol. Chem. 248, 1631-1633, 1973  
A>Title: The primary structure of human plasminogen. II. The histidine loop of human plasminogen  
A:Reference number: A92125; MUID:73149248; PMID:4694729  
A:Contents: annotation; active site  
R:Groskopf, W.R.; Summaria, L.; Robbins, K.C.  
J. Biol. Chem. 244, 3590-3597, 1969  
A>Title: Studies on the active center of human plasmin. Partial amino acid sequence of a human plasminogen  
A:Reference number: A92048; MUID:69234739; PMID:4420117  
A:Contents: annotation; active site  
R:Trexler, M.; Vali, Z.; Pathy, L.

J. Biol. Chem. 257, 7401-7406, 1982  
A>Title: Structure of the omega-aminocarboxylic acid-binding sites of human plasminogen  
A:Reference number: A92382; MUID:82213905; PMID:6919539  
A:Contents: annotation; omega-aminocarboxylic acid binding sites  
R:Vali, Z.; Pathy, L.  
J. Biol. Chem. 259, 13690-13694, 1984  
A>Title: The fibrin-binding site of human plasminogen. Arginines 32 and 34 are essential for fibrin binding  
A:Reference number: A92458; MUID:85054794; PMID:6094526  
A:Contents: annotation; fibrin binding site; omega-aminocarboxylic acid binding site  
R:Cao, Y.; Ji, R.W.; Davidson, D.; Schaller, J.; Marti, D.; Soemmel, S.; Mucca, P.; S.C.  
J. Biol. Chem. 271, 29461-29467, 1996  
A>Title: Kringle domains of human angiotensinogen. Characterization of the anti-plasminogen  
A:Reference number: A58811; MUID:97067211; PMID:8910613  
A:Contents: annotation  
R:Lijnen, H.R.; Ugar, F.; Bini, A.; Collen, D.  
Biochemistry 37, 4699-4702, 1998  
A>Title: Generation of an angiotensin-like fragment from plasminogen by stromelysin  
A:Reference number: A58812; MUID:9548733; PMID:9548733  
A:Contents: annotation  
R:Tulinsky, A.; Mulichak, A.M.  
Submitted to the Brookhaven Protein Data Bank, July 1991  
A:Reference number: A51341; PDB:1PK4  
A:Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 376-454  
R:Tulinsky, A.; Wu, T.P.  
Submitted to the Brookhaven Protein Data Bank, July 1991  
A:Reference number: A51488; PDB:2PK4  
A:Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 375-45  
R:Wu, T.P.; Tulinsky, A.  
Submitted to the Brookhaven Protein Data Bank, August 1993  
A:Reference number: A51911; PDB:1PKR  
A:Contents: annotation; X-ray crystallography, 2.48 angstroms, residues 102-18  
R:Padmanabhan, K.; Tulinsky, A.  
Submitted to the Brookhaven Protein Data Bank, April 1994  
A:Reference number: A52408; PDB:1PMK  
A:Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 377-45  
R:Tulinsky, A.; Mathews, I.I.  
Submitted to the Brookhaven Protein Data Bank, December 1995  
A:Reference number: A65244; PDB:1CEA  
A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181  
R:Tulinsky, A.; Mathews, I.I.  
Submitted to the Brookhaven Protein Data Bank, December 1995  
A:Reference number: A65245; PDB:1CEB  
A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181  
R:Mulichak, A.M.; Tulinsky, A.; Ravichandran, K.G.  
Biochemistry 30, 10576-10588, 1991  
A>Title: Crystal and molecular structure of human plasminogen kringle 4 refined at 1.9  
A:Reference number: A58819; MUID:92031502; PMID:1657148  
A:Contents: annotation  
R:Wu, T.P.; Padmanabhan, K.; Tulinsky, A.; Mulichak, A.M.  
Biochemistry 30, 10589-10594, 1991  
A>Title: The refined structure of the epsilon-aminocaproic acid complex of human plasminogen  
A:Reference number: A58818; MUID:92031503; PMID:1657149  
A:Contents: annotation  
R:de Vos, A.M.; Ultsch, M.H.; Kelley, R.F.; Padmanabhan, K.; Tulinsky, A.; Wesbrook, C.  
Biochemistry 31, 270-279, 1992  
A>Title: Crystal structure of the kringle 2 domain of tissue plasminogen activator  
A:Reference number: A39483; MUID:92118803; PMID:1310033  
A:Contents: annotation; X-ray crystallography, 2.4 angstroms  
R:Stec, B.; Teeter, M.M.; Whitlow, M.; Yamano, A.  
Submitted to the Brookhaven Protein Data Bank, June 1995  
A:Reference number: A65980; PDB:1KRN  
A:Contents: annotation; X-ray crystallography, 1.67 angstroms, residues 376-45  
R:Rejante, M.; Llinas, M.  
Submitted to the Brookhaven Protein Data Bank, August 1996  
A:Reference number: A65803; PDB:1HPJ  
A:Contents: annotation; conformation by (1)H-NMR, residues 103-181  
R:Rejante, M.; Llinas, M.  
Submitted to the Brookhaven Protein Data Bank, August 1996  
A:Reference number: A65804; PDB:1HPK  
A:Contents: annotation; conformation by (1)H-NMR, residues 103-181  
R:Rejante, M.R.; Llinas, M.  
Eur. J. Biochem. 221, 927-937, 1994  
A>Title: (1)H-NMR assignments and secondary structure of human plasminogen kringle 4



A;Reference number: S43645; MUID:94237157; PMID:8181475  
A;Content: annotation; conformation by (1)H-NMR, residues 96-184  
R;Refante, M.R.; Llinas, M.  
Eur. J. Biochem. 221, 939-949, 1994  
A;Title: Solution structure of the epsilon-aminohexanoic acid complex of human plasminogen  
A;Reference number: A58817; MUID:94237158; PMID:8181476  
A;Content: annotation; conformation by (1)H-NMR  
A;Comment: Plasminogen is synthesized by the kidney and is present in plasma and many other tissues.  
C;Comment: Plasminogen is converted to plasmin by plasminogen activators (see PIR:KUHUA and PIR:FGHUGB).  
C;Genetics:  
A;Gene: GDB:PLG  
A;Cross-references: GDB:119498; OMIM:173350  
A;Map position: 6q26-6q27  
A;Introns: 17/1; 62/2; 98/1; 136/2; 183/1; 223/2; 263/1; 317/2; 366/1; 419/2; 480/1; 529/1  
C;Function:  
A;Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a variety of the walls of the graafian follicle; also activates the urokinase-type plasminogen activator  
A;Pathway: fibrinolysis  
C;Superfamily: plasmin; kringie homology; plasminogen-related protein precursor homology  
C;Keywords: angiogenesis inhibitor; blood; duplication; fibrinolysis; glycoprotein; hydrolysis  
F;1-96/Domain: plasminogen-related protein precursor homology <PLPH>  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-810/Product: plasminogen #status experimental <PRO>  
F;20-96/Domain: activation peptide #status experimental <APT>  
F;79-466/Product: angiotensin #status experimental <AST>  
F;97-580,581-810/Product: plasmin #status experimental <NAT>  
F;97-580/Domain: plasmin chain A #status experimental <NAT>  
F;103-181/Domain: plasminogen homology <KR1>  
F;185-262/Domain: kringie homology <KR2>  
F;275-352/Domain: kringie homology <KR3>  
F;377-454/Domain: kringie homology <KR4>  
F;481-560/Domain: kringie homology <KR5>  
F;550-580,581-810/Product: microplasmin #status experimental <MWT>  
  
Query Match 19.4%; Score 154; DB 1; Length 810;  
Best Local Similarity 28.0%; Pred. No. 2.5e-06;  
Matches 45; Conservative 20; Mismatches 50; Indels 46; Gaps 10;  
  
QY 5 HQVPSNDCCLNGGTCVSNKYFSN-----IHWKNCPC-KKFGGQHCET-----DKS----- 47  
DB 308 NRTPEPCKN-----LDENYCRNPDGKRAPWCHTNSQVRWEYCKIPSCDSSPVSTEQLA 363  
QY 48 -----KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOYTHAHR---SDALQ 93  
DB 364 PTAPPELTPVQDCYHGDGSGYRGTSSTTTTGKQCQSWSS-----NTPHRHQKTPENYFN 418  
QY 94 LGLGKXNYCRNPDNRRPWCYVQGLKPLV--QECMVHDC A 132  
DB 419 AGL-TMNYCRNPDADKGWCF--TDFSVRWEYCNLKKS 455  
  
RESULT 34  
146260  
Plasmin (EC 3.4.21.7) precursor - western European hedgehog  
C;Species: Erinaceus europaeus (western European hedgehog)  
C;Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 16-Jul-1999  
C;Accession: I46260  
R;Lawn, R.M.; Boonmark, N.W.; Schwartz, K.; Lindahl, G.E.; Wade, D.P.; Byrne, C.D.; Fong, J. Biol. Chem. 270, 24004-24009, 1995  
A;Title: The recurring evolution of Lp(a): Insights from cloning of hedgehog apolipoprotein B  
A;Reference number: I46259; MUID:96025778; PMID:7592557  
A;Accession: I46260  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-810 <LAW>  
C;Cross-references: EMBL:U33171; NID:q1046360; PID:q1046361  
C;Superfamily: plasmin; kringie homology; plasminogen-related protein precursor homology

C;Keywords: hydrolase; serine proteinase  
F;1-96/Domain: plasminogen-related protein precursor homology <PLPH>  
F;103-181/Domain: kringie homology <KR1>  
F;185-262/Domain: kringie homology <KR2>  
F;275-352/Domain: kringie homology <KR3>  
F;379-456/Domain: kringie homology <KR4>  
F;482-561/Domain: kringie homology <KR5>  
F;582-803/Domain: trypsin homology <TRY>  
  
Query Match 19.4%; Score 153.5; DB 2; Length 810;  
Best Local Similarity 27.3%; Pred. No. 2.8e-06;  
Matches 39; Conservative 12; Mismatches 47; Indels 45; Gaps 6  
  
QY 6 QVPSNDCCLN-----GGTCVSNKYFSNIHW--CNCPKKFGGQHCETDKSKT- 49  
DB 309 RTPENYCKNLNENYCRNPDGEPAPWCFTTN--SSVRWEPCKIP-----DCVSSASETE 360  
QY 50 -----CYESGNGHFYRGKASTDTMGRPCLPWN SATVLQOYTHAHRSDA 91  
DB 361 HSDAPVTPPEPTQVVOEQYCGNGQTYRGTSSTTTTGKQCQPTSMRPHRSKTPENYD 420  
QY 92 LQGLGKXNYCRNPDNRRPWCY 114  
DB 421 ADLTM--NYCRNPDGDKGPWCY 440  
  
RESULT 35  
B30848  
Plasmin (EC 3.4.21.7) precursor - rhesus macaque  
C;Species: Macaca mulatta (rhesus macaque)  
C;Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 22-Jun-1999  
C;Accession: B32869; B30848  
R;Tomlinson, J.E.; McLean, J.W.; Lawn, R.M. J. Biol. Chem. 264, 5957-5965, 1989  
A;Title: Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of synthesis.  
A;Reference number: A32869; MUID:89174660; PMID:2925643  
A;Accession: B32869  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-810 <TOM>  
A;Cross-references: GB:J04697; NID:9342272; PID:AAA36901.1; PID:9342273  
C;Superfamily: plasmin; kringie homology; plasminogen-related protein precursor homology  
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringie; serine proteinase  
F;1-96/Domain: plasminogen-related protein precursor homology <PLPH>  
F;1-9/Domain: signal sequence #status predicted <SIG>  
F;103-181/Domain: kringie homology <KR1>  
F;185-262/Domain: kringie homology <KR2>  
F;275-352/Domain: kringie homology <KR3>  
F;377-454/Domain: kringie homology <KR4>  
F;481-560/Domain: kringie homology <KR5>  
F;581-803/Domain: trypsin homology <TRY>  
F;49-73,53-61,103-181,124-164,152-176,185-262,188-316,206-245,234-257,275-352,362,665,760/Active site: His, Asp, Ser #status predicted  
F;622,665,760/Active site: His, Asp, Ser #status predicted  
  
Query Match 19.2%; Score 152.5; DB 2; Length 810;  
Best Local Similarity 31.9%; Pred. No. 3.4e-06;  
Matches 45; Conservative 10; Mismatches 61; Indels 25; Gaps 6  
  
QY 10 NCDCLNGGTCVSNKYFSNIHWKNCPCKKFGGQ-----HCEIDKSKTCYEGNGH 56  
DB 428 NPDADKGWCFWCTTDPDSVRWEYCNLKKSCTGSGVAAPPPVAQLPDAETPSEDCWFGNGK 487  
QY 57 FYRGKASTDTMGRPCLPWN SATVLQOYTHAHR-----SDALQGLGKXNYCRNPD-NRRRP 111  
DB 488 GYRGKKAITVTGTGTCQEWAA-----QEPHSHRIETPTNPRAGLEK-NYCRNPDGVDGVP 541  
QY 112 WCYVQVGLKPLVQECMVHDC A 132  
DB 542 WCYT-TNPRKLFYDCVDPQCA 561  
  
RESULT 36



A40522  
plasma (EC 3.4.21.7) precursor - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 28-Feb-1992 #sequence\_revision 17-Apr-1993 #text\_change 16-Jul-1999  
C:Accession: A40522  
R:Kanalas J.J.; Makker S.P.  
J. Biol. Chem. 266, 10825-10829, 1991  
A:Title: Identification of the rat Heymann nephritis autoantigen (GP330) as a receptor  
A:Reference number: A40522; MUID:91250378; PMID:1645711  
A:Accession: A40522  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-169 <KAN>  
A:Cross-references: GB:N62832; NID:G206215; PIDN:AAA41884.1; PID:G554488  
A>Note: the authors translated the codon TCT for residue 76 as Ala  
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology  
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F:34-112/Domain: kringle homology <KRG>  
F:34-112,55-95,83-107/Disulfide bonds: #status predicted

Query Match 19.2%; Score 152; DB 2; Length 169;  
Best Local Similarity 32.2%; Pred. No. 1e-06;  
Matches 39; Conservative 16; Mismatches 30; Indels 36; Gaps 8;

QY 40 QHCEI-----DKS-----KTCYEKGHFYRGKASTDTMGRCLPWN SAT 78  
:  
Db 3 EYCEIPCGSSVSPDQSDSVLPQTVPVQECYGNGSKSYRGTSITNTGKCQSW---- 58

QY 79 VLQOTTYAHRS DALQL---GLGKNICYCRNDN--RRPWCVYVQVLKLIV--QECMVHDCA 132  
:  
Db 59 -VSMTPHSKSTPANFPDSGL-EMNYCRNPNDORGPCFT---TDPVSVMWEYNLKRCS 113

QY 133 D 133  
Db 114 E 114

RESULT 37  
B45082  
neurotrophic receptor ror2 precursor - human  
N:Contains: protein-tyrosine kinase (EC 2.7.1.112)  
C:Species: Homo sapiens (man)  
C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 19-May-2000  
C:Accession: B45082  
R:Masiakowski P.; Carroll R.D.  
J. Biol. Chem. 267, 26181-26190, 1992  
A:Title: A novel family of cell surface receptors with tyrosine kinase-like domain.  
A:Reference number: A45082; MUID:93100347; PMID:1334494  
A:Accession: B45082  
A:Molecule type: mRNA  
A:Residues: 1-943 <MAS>  
A:Cross-references: GB:M37639; NID:G337466; PIDN:AAA60276.1; PID:G337467  
A>Note: sequence extracted from NCBI backbone (NCBIP:120918)  
C:Genetics:  
A:Gene: GDB:NTRKR2  
A:Cross-references: GDB:I36454  
A:Map position: 6p21-6p21  
C:Superfamily: neurotrophic receptor ror; immunoglobulin homology; kringle homology; pro-  
C:Keywords: ATP; glycoprotein; kringle; phosphotransferase; transmembrane protein; tyro-  
F:1-27/Domain: signal sequence #status predicted <SIG>  
F:28-93/Product: neurotrophic receptor ror2 #status predicted <MAT>  
F:76-137/Domain: immunoglobulin homology <IMW>  
F:318-394/Domain: kringle homology <KRG>  
F:418-428/Domain: transmembrane #status predicted <TMN>  
F:471-753/Domain: protein kinase homology <KIN>  
F:479-487/Region: protein kinase ATP-binding motif  
F:70,188,318/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 18.7%; Score 148; DB 2; Length 943;  
Best Local Similarity 34.3%; Pred. No. 9.9e-06;  
Matches 37; Conservative 12; Mismatches 41; Indels 18; Gaps 6;

QY 32 NC-----PKFFGGCHCIDKSKTCYEKGHFYRGKASTDTMGRCLPWN SATVLQOOTH 85

DB 299 NCMTIGIPAEELGRVH-----OCYNGSGMDVRGTASTTKSGHQCOFPW--ALQHPSHA 349

QY 86 AHRSDALQLGLGKKNCRNPDRR-PWCYVQVLKELVQECMVHDCA 132

DB 350 LSSITDFPELG-GGHAYCNRPGCGEGPWCFTQ-NKNVMELCDVPSCS 395

RESULT 38  
PLMS

plasmin (EC 3.4.21.7) precursor - mouse  
N;Contains: angiotatin; plasminogen  
C;Species: Mus musculus (house mouse)  
C;Date: 20-Sep-1991 #sequence revision 01-Nov-1996 #text\_change 18-Jun-1999  
C;Accession: A38514; S48202; S48203  
R;Begen, S.J.F.; Bell, S.M.; Schaefer, L.A.; Elliott, R.W.  
Genomics 8, 49-61, 1990  
A;Title: Characterization of the cDNA coding for mouse plasminogen and localization of  
A;Reference number: A38514; MUID:91184812; PMID:2081600  
A;Accession: A38514  
A;Molecule type: mRNA  
A;Residues: 1-812 <DEG>  
A;Cross-references: GB|J04766; NID:G200402; PIDN:AAA50168.1; PID:G200403  
R;Lijnen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.  
Eur. J. Biochem. 224, 863-871, 1994  
A;Title: Characterization of the murine plasma fibrinolytic system.  
A;Reference number: S48202; MUID:95010076; PMID:7523120  
A;Accession: S48202  
A;Molecule type: Protein  
A;Residues: 20-25 <LDJ>  
A;Accession: S48203  
A;Molecule type: Protein  
A;Residues: 22-27 <LI2>  
C;Comment: Plasminogen is synthesized by the kidney and is present in plasma a  
C;Comment: Plasminogen is converted into plasmin by plasminogen activators, b  
mediately after dissociation from the clot. In the presence of the inhibitor,  
e inhibitor, the activation involves also removal of the activation peptide.  
C;Comment: Stromelysin 1 (see PIR:KMSS1) acts on plasminogen to produce angio  
tef in treating solid tumors.

C;Function:

A;Description: dissolves the fibrin of blood clots; acts as a proteolytic fact  
ns the walls of the graafian follicle; also activates the urokinase-type plas  
A;Pathway: fibrinolysis  
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precurs  
C;Keywords: angiogenesis inhibitor; blood; duplication; fibrinolysis; glycopro  
F;1-96/Domain: plasminogen-related protein precursor homology <PLPH>  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-812/Product: plasminogen #status predicted <PRO>  
F;20-96/Domain: activation peptide #status predicted <APT>  
F;79-46/Product: angiotatin #status predicted <AST>  
F;97-581,582-812/Product: plasmin #status predicted <MAT>  
F;97-581/Domain: chain A #status predicted <ACH>  
F;103-181/Domain: kringle homology <KR1>  
F;185-262/Domain: kringle homology <KR2>  
F;275-352/Domain: kringle homology <KR3>  
F;377-454/Domain: kringle homology <KR4>  
F;481-560/Domain: kringle homology <KR5>  
F;582-805/Domain: trypsin homology <TRY>  
F;49-73,53-61,103-181,124-164,152-176,185-262,188-316,206-245,234-257,275-352,  
onds: #status predicted  
F;78-79/Cleavage site: Glu-Asn (stromelysin 1) #status predicted  
F;136,308/Binding site: carbohydrate (asn) (covalent) #status predicted  
F;466-467/Cleavage site: Thr-Val (stromelysin 1) #status predicted  
F;581-582/Cleavage site: Arg-Val (plasminogen activator) #status experimental  
F;624,667,762/Active site: His, Asp, Ser #status predicted

Query Match 18.5%; Score 147; DB 1; Length 812;  
Best Local Similarity 26.4%; Pred. No. 1.1e-05;  
Matches 42; Conservative 21; Mismatches 56; Indels 40; Gaps 9

QY 5 HQVPSNCDCNLNGTCVSNKYFSN-----IHWCK-NCPKKFGGQHCRI-----DKS--- 47

Db 308 NRTPTNPFCKN-----LEENYCRNPDGHTAPWCYTTDSQLRWEYCEIPSCSSASPDQSDS 363  
Qy 48 -----KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGL 96  
Db 364 SVPPEQTPVVOEQYQSGQSYRGCTSTTTTGKKQCSW--AAMPFPHRSKTPENFFDAGL 421  
Qy 97 GKHYCRNPNRRPWCYVQGLKPLV--QECMVHDCAD 133  
Db 422 -EMNYCRNPDGDKGWCYCT---TDPVSRWEYCNLKRCE 456

## RESULT 39

T18840  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T18840  
R:Berks, M.  
submitted to the EMBL Data Library, August 1994  
A:Reference number: Z19029  
A:Accession: T18840  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-806 <WIL>  
A:Cross-references: EMBL:Z35595; PIDN:CAA84639.1; GSPDB:GN00020; CESP:CO1G6.8  
A:Experimental source: clone CO1G6  
C:Genetics:  
A:Map position: 2  
A:Introns: 36/3; 170/3; 217/3; 636/3; 760/1  
Query Match 18.2%; Score 144.5; DB 2; Length 806;  
Best Local Similarity 25.5%; Pred. No. 1.8e-05;  
Matches 42; Conservative 25; Mismatches 55; Indels 43; Gaps 9;

Qy 1 SNEHQVPS---NDCINGTCVSNKYFSNTH-----WCNCPKFGGQHC--- 42  
Db 152 SDSNNTIVSICKHDCDDVQNDSPSELAAQHVLGTPKALFPLCSLSTNSCIPVM 211  
Qy 43 -----EIDK---SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHR 89  
Db 212 STALQSSPAEVNVRGHLTWCVNSGTQYEGTVAQTSKGKQCAPWIDST--SRDPNVHRF 269  
Qy 90 DALQLGLGKHYCRNPNRR--RPMCVVQVGLKPLVQEC--CMVHDC 131  
Db 270 PEL---MNSKNYCRNPGGKKRPMWCY-----SKPMGQGEYCDVPQC 307

## RESULT 40

A47136  
macrophage-stimulating protein 1 precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 03-May-1994 #sequence\_revision 14-Nov-1997 #text\_change 18-Jun-1999  
C:Accession: A40331; B40331; A47136; A61395  
R:Han, S.; Stuart, L.A.; Degen, S.J.F.  
Biochemistry 30, 9768-9780, 1991  
A:Title: Characterization of the DNFI5S2 locus on human chromosome 3: identification of  
A:Reference number: A40331; MUID:920202016; PMID:1655021  
A:Accession: A40331  
A:Molecule type: DNA  
A:Residues: 1-711 <HA1>  
A:Cross-references: GB:M74179  
A:Accession: B40331  
A:Molecule type: mRNA  
A:Residues: 1-711 <HA2>  
A:Cross-references: GB:M74178; NID:G183976; PIDN:AAA50165.1; PID:G183977  
R:Yoshimura, T.; Yuhki, N.; Wang, M.H.; Skeel, A.; Leonard, E.J.  
J. Biol. Chem. 268, 15461-15469, 1993  
A:Title: Cloning, sequencing, and expression of human macrophage stimulating protein (MS  
A:Reference number: A47136; MUID:93340141; PMID:8393443  
A:Accession: A47136  
A:Molecule type: mRNA  
A:Residues: 1-12, 'C', 14-622, 'F', 624-711 <YOS>

A:Cross-references: GB:L11924; NID:G398037; PIDN:AAA59872.1; PID:G398038  
A:Note: authors translated the codon TTT for residue 623 as Leu; parts of this  
R:Skeel, A.; Yoshimura, T.; Showalter, S.D.; Tanaka, S.; Appella, E.; Leonard, E.  
J. Exp. Med. 173, 1227-1234, 1991  
A:Title: Macrophage stimulating protein: purification, partial amino acid sequ  
A:Reference number: A61395; MUID:91217635; PMID:1827141  
A:Accession: A61395  
A:Molecule type: protein  
A:Residues: 230-247; 288-291, 'E', 293-295, 'X', 297-301, 'X', 303, 'E', 305, 'EX', 308-3  
A:Experimental source: plasma  
C:Genetics:  
A:Gene: GDB:MST1; D3F15S2; DNFI5S2; HGFL  
A:Cross-references: GDB:128833; OMIM:142408  
A:Map position: 3p21-3p21.3  
C:Complex: disulfide-bonded heterodimer of chains derived from the same precu  
C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology  
C:Keywords: duplication; glycoprotein; growth factor; kringle; plasma  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-483/Domain: alpha chain #status predicted <ACH>  
F:19-483/Domain: kringle homology <KR1>  
F:110-186/Domain: kringle homology <KR2>  
F:191-268/Domain: kringle homology <KR3>  
F:283-361/Domain: kringle homology <KR4>  
F:370-448/Domain: kringle homology <KR4>  
F:484-711/Domain: beta chain #status predicted <BCH>  
F:484-704/Domain: trypsin homology <TRY>  
F:56-78, 60-66, 110-186, 131-169, 157-181, 191-268, 212-251, 240-263, 283-361, 304-343,  
F:72, 236, 615/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 18.2%; Score 144; DB 1; Length 711;  
Best Local Similarity 27.0%; Pred. No. 1.8e-05;  
Matches 33; Conservative 19; Mismatches 44; Indels 26; Gaps 6  
Qy 30 WC-NCPKFGGQHCIDK-----SKTCYEGNGHFYRGKASTDTMGRPCLPWNS 76  
Db 250 WCYTTDPQIEREFCDLPRGSEAPQREATTVCSPFGKGEYRGATNTTAGVPCORWDA 309  
Qy 77 ATVLQOQTYHAHRSDALQLGLG--KHNYCRNPNRRPWCYVQVGLKPLVQEC---MVHDC 131  
Db 310 -----QIPHOHRTPEKYACKDLRFNCFNPDGSEAPWCFT---LPGHRAAPCYQIRRC 361  
Qy 132 AD 133  
Db 362 TD 363

## RESULT 41

A61545  
plasmin (EC 3.4.21.7) precursor - horse (fragments)  
N:Alternate names: plasminogen  
N:Contains: miniplasminogen  
C:Species: Equus caballus (domestic horse)  
C:Date: 28-Oct-1994 #sequence\_revision 01-Nov-1996 #text\_change 18-Jul-1997  
C:Accession: A61545; S17527  
R:Schaller, J.; Rickli, E.E.  
Enzyme 40, 63-69, 1988  
A:Title: Structural aspects of the plasminogen of various species.  
A:Reference number: A61545; MUID:89005015; PMID:3168975  
A:Accession: A61545  
A:Molecule type: protein  
A:Residues: 1-33, 34-117 <SCH>  
R:Schaller, J.; Straub, C.; Kaempfer, U.; Rickli, E.E.  
Protein Seq. Data Anal. 4, 69-74, 1991  
A:Title: Complete amino acid sequence of equine miniplasminogen.  
A:Reference number: S17527; MUID:92052077; PMID:1946332  
A:Accession: S17527  
A:Molecule type: protein  
A:Residues: 118-455 <SC2>  
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor  
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine pro  
F:1-33, 34-117, 118-455/Product: plasminogen (fragments) #status experimental <F  
F:1-33/Domain: activation peptide (fragment) #status experimental <APT>  
F:34-117, 118-225, 226-455/Product: plasmin (fragments) #status experimental <MA

F:37-114/Domain: kringle homology #K4>  
F:118-455/Product: miniplasminogen #K5>  
F:126-205/Domain: kringle homology #K5>  
F:226-455/Domain: plasmin chain B #status experimental <BCH>  
F:226-448/Domain: trypsin homology <TRY>  
F:267,310,405/Active site: His, Asp, Ser #status predicted

Query Match 17.8%; Score 141.5; DB 2; Length 455;  
Best Local Similarity 33.7%; Pred. No. 2.1e-05;  
Matches 33; Conservative 15; Mismatches 37; Indels 13; Gaps 5;  
QY 40 QHCEIDSKTCYEGNGHYRKGASTDTMGRCLPNSATV--LQOTYHRSALQLGLG 97  
DB 28 BECSA-KVQDCYQDKGSEYRTSSITVTGKCKQSWMTPHWQKTPKYPNADLTW--- 83

QY 98 KHNTRNPDNRPRPCVQVGLKPLV--QECWHDICAD 133  
DB 84 --NYCRNPDGKGPWCYT--TDSVRWEFCNLKQSE 116

## RESULT 42

JH0579

hepatocyte growth factor precursor [validated] - human

N:Alternate names: hepatoprotein A; scatter factor

C:Species: Homo sapiens (man)

C&gt;Date: 17-Aug-1992 #sequence revision 17-Aug-1992 #text change 08-Dec-2000

C/Accession: JH0579; J00333; A41140; B36677; A33512; A39006; PH0114; A37796; S06

R:Seiki, T.; Hagiya, M.; Shimomishi, M.; Nakamura, T.; Shimizu, S.

Gene 102, 213-219, 1991

A:Title: Organization of the human hepatocyte growth factor-encoding gene.

A:Reference number: JH0579; MUID:91340155; PMID:1831432

A:Accession: JH0579

A:Molecule type: DNA

A:Residues: 1-728 &lt;SEK&gt;

A:Note: The authors translated the codon GAA for residue 662 as Gly

R:Seiki, T.; Hagiya, M.; Shimomishi, M.; Nakamura, T.; Shimizu, S.

submitted to JPIB, March 1991

A:Description: Organization of the human hepatocyte growth factor-encoding gene.

A:Reference number: JH0579

A:Accession: JH0579

A:Molecule type: DNA

A:Residues: 1-481, RT, 484-728 &lt;SE2&gt;

R:Weidner, K.M.; Arakaki, N.; Hartmann, G.; Vandekerckhove, J.; Weingart, S.; Rieder, H.

Proc. Natl. Acad. Sci. U.S.A. 88, 7001-7005, 1991

A:Title: Evidence for the identity of human scatter factor and human hepatocyte growth factor

A:Reference number: A41140; MUID:91334393; PMID:1831266

A:Accession: A41140

A:Molecule type: mRNA

A:Residues: 1-728 &lt;WEI&gt;

A:Cross-references: GB:M73239; NID:9337935; PIDN:AAA64239.1; PID:G337936

R:Seiki, T.; Ihara, I.; Sugimura, A.; Shimomishi, M.; Nishizawa, T.; Asami, O.; Hagiya, M.

Biochem. Biophys. Res. Commun. 172, 321-327, 1990

A:Title: Isolation and expression of cDNA for different forms of hepatocyte growth factor

A:Reference number: A36677; MUID:91025062; PMID:2145836

A:Accession: B36677

A:Molecule type: mRNA

A:Residues: 1-728 &lt;SE3&gt;

A:Cross-references: GB:M60718; NID:9184031; PIDN:AAA52648.1; PID:G184032

A:Accession: A36677

A:Molecule type: mRNA

A:Residues: 1-161, 167-728 &lt;SE4&gt;

A:Experimental source: leukocyte

R:Miyaizawa, K.; Tsubouchi, H.; Naka, D.; Takahashi, K.; Okigaki, N.; Nakaya

Biochem. Biophys. Res. Commun. 163, 967-973, 1989

A:Title: Molecular cloning and sequence analysis of cDNA for human hepatocyte growth factor

A:Reference number: A33512; MUID:89392017; PMID:2528952

A:Accession: A33512

A&gt;Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-728 &lt;MIY&gt;

A:Cross-references: GB:M29145; NID:9184041; PIDN:AAA52650.1; PID:G306846

R:Rubin, J.S.; Chan, A.M.L.; Bottaro, D.P.; Burgess, W.H.; Taylor, W.G.; Cech, C.; H  
Proc. Natl. Acad. Sci. U.S.A. 88, 415-419, 1991  
A:Title: A broad-spectrum human lung fibroblast-derived mitogen is a variant of  
A:Reference number: A39006; MUID:91110540; PMID:1824873  
A:Accession: A39006  
A:Molecule type: mRNA  
A:Residues: 1-161, 167-728 <RUB>  
A:Cross-references: GB:M55379  
A:Experimental source: embryonic lung  
R:Yoshiyama, Y.; Arakaki, N.; Naka, D.; Takahashi, K.; Hirono, S.; Kondo, J.; Hayama  
Biochem. Biophys. Res. Commun. 175, 660-667, 1991  
A:Title: Identification of the N-terminal residue of the heavy chain of both n  
A:Reference number: PH0114; MUID:91207365; PMID:1826837  
A:Accession: PH0114  
A:Molecule type: protein  
A:Residues: 32-43; 53-58 <YOS>  
A:Experimental source: plasma  
R:Weidner, K.M.; Behrens, J.; Vandekerckhove, J.; Birchmeier, W.  
J. Cell Biol. 111, 2097-2108, 1990  
A:Title: Scatter factor: molecular characteristics and effect on the invasiveness  
A:Reference number: A37796; MUID:91035621; PMID:2148276  
A:Accession: A37796  
A:Molecule type: protein  
A:Residues: 86-91; 329-344; 356-363, 'XX', 366-370; 425-434; 442-447, 'X', 449-450; 543  
R:Nakamura, T.; Nishizawa, T.; Seki, T.; Shimomishi, M.; Sugimura, T.; Tas  
Nature 342, 440-443, 1989  
A:Title: Molecular cloning and expression of human hepatocyte growth factor.  
A:Reference number: S06794; MUID:90066676; PMID:2531289  
A:Accession: S06794  
A:Molecule type: mRNA  
A:Residues: 1-31, 'HK', 34-77, 'N', 79-292, 'V', 294-299, 'M', 301-316, 'A', 318-335, 'K', 337-386  
A:Cross-references: EMBL:X16323; NID:G32081; PIDN:CAA34387.1; PID:G32082  
A:Note: the authors translated the codon CAG for residue 727 as Glu  
A:Note: part of this sequence, including the amino end of both the alpha and b  
R:Hartmann, G.; Naldini, L.; Weidner, K.M.; Sachs, M.; Vigna, E.; Comoglio, P.; Birc  
Proc. Natl. Acad. Sci. U.S.A. 89, 11574-11578, 1992  
A:Title: A functional domain in the heavy chain of scatter factor/hepatocyte g  
A:Reference number: I59214; MUID:93087571; PMID:1280830  
A:Accession: I59214  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-288, 'ET' <HAR>  
A:Cross-references: GB:I02931; NID:9184033; PIDN:AAA52649.1; PID:G184034  
R:Miyaizawa, K.; Kitamura, A.; Naka, D.; Kitamura, N.  
Eur. J. Biochem. 197, 15-22, 1991  
A:Title: An alternatively processed mRNA generated from human hepatocyte growth factor  
A:Reference number: S15443; MUID:91200041; PMID:1826653  
A:Accession: S15443  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-288, 'ET' <MIY2>  
A:Cross-references: EMBL:X57574; NID:G32083; PIDN:CAA40802.1; PID:G32084  
R:Shima, N.; Nagao, M.; Ogaki, F.; Tsuda, E.; Murakami, A.; Higashio, K.  
Biochem. Biophys. Res. Commun. 180, 1151-1158, 1991  
A:Title: Tumor cytotoxic factor/hepatocyte growth factor from human fibroblast  
A:Reference number: I52253; MUID:92062058; PMID:1835383  
A:Accession: I52253  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 161-166 <SHI>  
A:Cross-references: GB:S62561; NID:G237996; PIDN:AAB20169.1; PID:G237997  
C:Genetics:  
A:Gene: GDB:HGF  
A:Cross-references: GDB:I127524; OMIM:142409  
A:Map position: 7q21.1-7q21.1  
A:Introns: 30/1; 85/2; 123/1; 161/2; 209/1; 249/2; 289/1; 347/2; 390/1; 424/2; 491/1;  
C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor  
C:Function:  
A:Description: stimulates mitosis of hepatocytes and other cells  
A:Note: does not have proteinase activity  
C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology  
C:Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringle; p

F:1-31/Domain: signal sequence #status predicted <SIG>  
F:32-494/Domain: hepatocyte growth factor #status experimental <MAT>  
F:128-206/Domain: kringles homology <KR1>  
F:211-286/Domain: kringles homology <KR2>  
F:305-383/Domain: kringles homology <KR3>  
F:391-469/Domain: kringles homology <KR4>  
F:495-728/Domain: beta chain #status experimental <BCH>  
F:495-716/Domain: tryptophan homology <TRY>  
F:32/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimental  
F:294, 402, 566, 653/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:487-604/Disulfide bonds: #status predicted

Query Match 17.8%; Score 141.5; DB 1; Length 728;  
Best Local Similarity 30.3%; Pred. No. 3.1e-05;  
Matches 44; Conservative 12; Mismatches 64; Indels 25; Gaps 10;

QY 2 NELHGVPSNCDG--LNGGTCVSNKYFNSIHWG-----NCPKFGGQ--HCIDSKSKTYE 52  
DB 335 HEHDTWTFNFKCKDLRENYC--RNPDSGSPWCFTTDPNIRVGYCSQIPNCDMSHGQDCYR 393  
QY 53 GNGHGYRKGASTDTMGRPCLPNSATVLOQTYVHAH---RSDALQLGLGKHNKYNCRNP--DNR 108  
DB 394 GNGKNTGMNLSQTRSLGTCSMWDKN---MEDLHRHIFWEPDASKL---NENYCRNPDDA 447  
QY 109 RBPWCYVQGLKPLV--QECMVHDC 131  
DB 448 HGPWCYTG---NPLIPWDYCPISRC 469

RESULT 43  
I51285  
hepatocyte growth factor/scatter factor - chicken (fragment)  
C/Species: Gallus gallus (chicken)  
C/Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 16-Jul-1999  
C/Accession: I51285  
R:Streit, A.; Stern, C.D.; Thery, C.; Ireland, G.W.; Aparicio, S.; Sharpe, M.J.; Gherard  
Development 121, 813-824, 1995  
A:Title: A role for HGF/SF in neural induction and its expression in Hensen's node during  
A:Reference number: I51285; MUID:95237013; PMID:7720585  
A:Accession: I51285  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-411 <STR>  
A:Cross-references: GB:S77480; NID:G998675; PID:G998676  
C/Superfamily: hepatocyte growth factor; kringles homology; trypsin homology  
F:124-197/Domain: kringles homology <KR1>  
F:202-279/Domain: kringles homology <KR2>  
F:296-374/Domain: kringles homology <KR3>

Query Match 17.8%; Score 141; DB 2; Length 411;  
Best Local Similarity 28.5%; Pred. No. 2.1e-05;  
Matches 37; Conservative 18; Mismatches 45; Indels 30; Gaps 7;

QY 30 WC-----NCPKFGG-QHCEID-----KSKTCYEGNGHFGKASTDTMGRPCLPW 74  
DB 261 WCYLDPNTPWEFCAIKTCDVGLNSTEAVETTTICQGGEGYRGTVNTIWSGIQCORW 320  
QY 75 NSATVLQQTYVHAH--SDALQLGLGKHNKYNCRNPDRRRPWCY-----VQVGLKPLVQRCM 127  
DB 321 DS-----QRPQHNTIPENFKCKDLRENYCRNPDSGSPWCFTTDPNIRVGYCSQIPKCD 375  
QY 128 V----HCADG 134  
DB 376 VSNEQDCYRG 385

RESULT 44  
I51283  
hepatocyte growth factor precursor - clawed frog  
N/Alternate names: hepatopoietin A; scatter factor  
C/Species: Xenopus sp. (clawed frog)  
C/Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 18-Jun-1999

C/Accession: I51283  
R:Nakamura, H.; Tashiro, K.; Nakamura, T.; Shiohara, K.  
Mech. Dev. 49, 123-131, 1995  
A:Title: Molecular cloning of Xenopus HGF cDNA and its expression studies in Xopus e  
A:Reference number: I51283; MUID:95267690; PMID:7748783  
A:Accession: I51283  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-710 <NAX>  
A:Cross-references: GB:S77422; NID:G998932; PIDN:AA834354.1; PID:G998933  
A:Note: the authors' translation for residue 458 (Thr) is inconsistent with the nucleo  
C/Complex: disulfide-bonded heterodimer of chains derived from the same precursor  
C/Function:  
A:Description: stimulates mitosis of hepatocytes and other cells  
A:Note: Does not have proteinase activity  
C/Superfamily: hepatocyte growth factor; kringles homology; trypsin homology  
C/Keywords: duplication; glycoprotein; growth factor; heterodimer; kringles  
F:42-477/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>  
F:42-477/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>  
F:115-193/Domain: kringles homology <KR1>  
F:198-275/Domain: kringles homology <KR2>  
F:289-367/Domain: kringles homology <KR3>  
F:375-453/Domain: kringles homology <KR4>  
F:478-709/Domain: hepatocyte growth factor beta chain #status predicted <BCH>  
F:478-700/Domain: tryptophan homology <TRY>  
F:52, 128, 281, 322, 379, 550, 637, 666/Binding site: carbohydrate (Asn) (covalent) #  
F:470-588/Disulfide bonds: #status predicted

Query Match 17.8%; Score 141; DB 1; Length 710;  
Best Local Similarity 29.2%; Pred. No. 3.4e-05;  
Matches 42; Conservative 12; Mismatches 64; Indels 26; Gaps 9

QY 4 LHQ-VPSNCDG--LNGGTCVSNKYFNSIHWG-----NCPKFGGQHCIDSKSKTC 50  
DB 320 LHNFTENYKCDLSENYC--RNPDSGSPWCFTTDPNIRIGHCSQI---KKQASNQEC 375  
QY 51 YEGNGHFGKASTDTMGRPCLPNSATVLOQTYVHAHSDALQLGLGKHNKYNCRNP--RR 109  
DB 376 YGNGSTYKGLTSLRTRFLPCSMWEKNL---QDLKRTFTNEPDSVLSIQRNCPNDAAH 432  
QY 110 RPYCYVQGLKPLV--QECMVHDC 131  
DB 433 GPCWYTD---DPFVWDYCPISRC 453

RESULT 45  
A40332  
macrophage-stimulating protein 1 precursor - mouse  
N/Alternate names: hepatocyte growth factor-like protein  
C/Species: Mus musculus (house mouse)  
C/Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 18-Jun-1999  
C/Accession: A40332; B40332  
R:Degen, S.J.F.; Stuart, L.A.; Han, S.; Jamison, C.S.  
Biochemistry 30, 9781-9791, 1991  
A:Title: Characterization of the mouse cDNA and gene coding for a hepatocyte g  
A:Reference number: A40332; MUID:92002017; PMID:1632957  
A:Accession: A40332  
A:Molecule type: DNA  
A:Residues: 1-716 <DEG>  
A:Cross-references: GB:M74180; NID:G193831; PIDN:AAA50166.1; PID:G193832  
A:Accession: B40332  
A:Molecule type: mRNA  
A:Residues: 1-18, 'P', 20-715 <DEG2>  
A:Cross-references: GB:M74181; NID:G193833; PIDN:AAA50167.1; PID:G193834  
C/Genetics:  
A:Introns: 18/1; 67/2; 105/1; 143/2; 189/1; 229/2; 269/1; 334/2; 378/1; 412/2;  
C/Complex: disulfide-bonded heterodimer of chains derived from the same precursor  
C/Superfamily: hepatocyte growth factor; kringles homology; trypsin homology  
C/Keywords: duplication; glycoprotein; growth factor; kringles  
F:1-31/Domain: signal sequence #status predicted <SIG>  
F:19-488, 489-716/Domain: macrophage-stimulating protein 1 #status experimental  
F:19-483/Domain: alpha chain #status experimental <ACH>  
F:110-186/Domain: kringles homology <KR1>

F:191-268/Domain: kringle homology <KR2>  
F:292-370/Domain: kringle homology <KR3>  
F:379-457/Domain: kringle homology <KR4>  
F:484-711/Domain: beta chain #status experimental <BCH>  
F:499-709/Domain: trypsin homology <TRY>  
F:712,173,305,620/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match 17.5%; Score 139; DB 1; Length 716;  
Best Local Similarity 39.6%; Pred. No. 5.1e-05;  
Matches 38; Conservative 4; Mismatches 32; Indels 22; Gaps 7;

QY 30 WC---NCPKFGGQHCEIDKSKTYEG-----NGHFYRGKASTDTMGRPCLPWN SATVLQ 81  
DB 168 WCYTTRNSVRP--QSCGI--KTCEAVCLVNCDEYRGEVDVTESGRCQRWD-----L 217  
QY 82 QTYHAH---RSDALQLGLGKHYCRNPNNRRPWCY 114  
DB 218 QHPSHHPQPEKFLDKDL-KQNYCRNPDGSRPWCY 252

RESULT 46  
A35644  
hepatocyte growth factor precursor - rat  
N:Alternate names: hepatopietin A; scatter factor  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 28-Sep-1990 #sequence\_revision 18-Nov-1992 #text\_change 21-Jul-2000  
C:Accession: A35644; S13211  
R:Tachiro, K.; Hagiya, M.; Nishizawa, T.; Seki, T.; Shimonishi, M.; Shimizu, S.; Nakamura, Proc. Natl. Acad. Sci. U.S.A. 87, 3200-3204, 1990  
A>Title: Deduced primary structure of rat hepatocyte growth factor and expression of the  
A:Reference number: A35644; MUID:9022197; PMID:2135229  
A:Accession: A35644  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-728 <TAS>  
A:Cross-references: GB:D90102; GB:M32987; NID:G220766; PIDN:BAAL4133.1; PID:G220767  
A>Note: The authors translated the codon GAG for residue 70 as Gln, GAC for residue 417  
R:Okajima, A.; Miyazawa, K.; Kitamura, N.  
Eur. J. Biochem. 193, 375-381, 1990  
A>Title: Primary structure of rat hepatocyte growth factor and induction of its mRNA dur  
A:Reference number: S13211; MUID:91031482; PMID:2146117  
A:Accession: S13211  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-728 <OKA>  
A:Cross-references: EMBL:X54400; NID:G56353; PIDN:CAA38266.1; PID:G4539554  
C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor  
C:Function:  
A:Note: does not have proteinase activity  
C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology  
C:Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringle; pyn  
F:1-32/Domain: signal sequence #status predicted <SIG>  
F:56-495/Product: hepatocyte growth factor #status predicted <MAT>  
F:56-495/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>  
F:129-207/Domain: kringle homology <KR1>  
F:212-289/Domain: kringle homology <KR2>  
F:306-384/Domain: kringle homology <KR3>  
F:392-470/Domain: kringle homology <KR4>  
F:496-728/Domain: hepatocyte growth factor beta chain #status predicted <BCH>  
F:496-719/Domain: trypsin homology <TRY>  
F:733/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted  
F:295,403,569,656/Binding site: carboxylate (Asn) (covalent) #status predicted  
F:488-607/Disulfide bonds: #status predicted

Query Match 17.5%; Score 139; DB 1; Length 728;  
Best Local Similarity 29.5%; Pred. No. 5.2e-05;  
Matches 44; Conservative 10; Mismatches 57; Indels 38; Gaps 11;

QY 5 HQV-PSNDC--LNGGTCVSNKYFQNIHWC-----NCPKFGGQHCEIDKSK 48  
DB 338 HDITPENFKDLRENYC-RNPDGAESPWCFTTDPNIRVGYCSQIPK-----CDVSSGQ 390

QY 49 TCYEGNGHYRGKASTDTMGRPCLPWN SATVLQ QTYHAH---RSDALQLGLGKHYCRNP 105  
DB 391 DCYRGNGKNGKNGNLKSTRSGLTCSNWDKN---MEDLHRHIFWEPDASKL---TKNYCRNP 444  
QY 106 -DNRRRPWCYQVQGLKPLV--QECMVHDC 131  
DB 445 DDDAGHPWCYTG--NPLVPWDYCPISRC 470

RESULT 47  
A60185  
hepatocyte growth factor precursor - mouse  
N:Alternate names: hepatopietin A; scatter factor  
C:Species: Mus musculus (house mouse)  
C>Date: 03-Mar-1993 #sequence\_revision 26-May-1994 #text\_change 16-Jun-2000  
C:Accession: JG2117; PC2064; A60185; S43416; S45521; S17173; S10966; I48758; J 231  
R:Sasaki, M.; Nishio, M.; Sasaki, T.; Enami, J.  
Biochem. Biophys. Res. Commun. 199, 772-779, 1994  
A>Title: Identification of mouse mammary fibroblast-derived mammary growth fac  
A:Reference number: JG2117; MUID:94183257; PMID:8135822  
A:Accession: JG2117  
A:Molecule type: mRNA  
A:Residues: 1-728 <SAS2>  
A:Cross-references: GB:D10212; NID:G220435; PIDN:BAAO1064.1; PID:G220436  
A:Experimental source: fibroblast, COS-1 cell  
A>Note: submitted to JIPID, May 1993  
A:Accession: PC2064  
A:Molecule type: protein  
A:Residues: 496-504 <SA2>  
R:Rosen, E.M.; Meromsky, L.; Setter, E.; Vinter, D.W.; Goldberg, I.D.  
Proc. Soc. Exp. Biol. Med. 195, 34-43, 1990  
A>Title: Purified scatter factor stimulates epithelial and vascular endothelia  
A:Reference number: A60185; MUID:90377927; PMID:2144630  
A:Accession: A60185  
A:Molecule type: protein  
A:Residues: X', 184-188, 'XX', 191-192, 'X', 194, 'XX', 197, 357-364, 'XX', 367, 375-377  
R:Liu, Y.; Michalopoulos, G.K.; Zarnegar, R.  
Biochim. Biophys. Acta 1216, 299-303, 1993  
A>Title: Molecular cloning and characterization of cDNA encoding mouse hepatoc  
A:Reference number: S43416; MUID:94060105; PMID:8241272  
A:Accession: S43416  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-728 <LIU>  
A:Cross-references: EMBL:X72307  
R:Liu, Y.  
submitted to the EMBL Data Library, May 1993  
A:Reference number: S45521  
A:Accession: S45521  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-563, 'H', 565-728 <LI2>  
R:Coffier, A.; Fellows, J.; Young, S.; Pappin, D.; Rahman, D.  
Biochem. J. 278, 35-41, 1991  
A>Title: Purification and characterization of biologically active scatter fact  
A:Reference number: S17173; MUID:91354223; PMID:1831975  
A:Accession: S17173  
A:Molecule type: protein  
A:Residues: 496-517, 'T', 519 <COF>  
R:Gherardi, E.; Stoker, M.  
Nature 346, 228, 1990  
A>Title: Hepatocytes and scatter factor.  
A:Reference number: S10966; MUID:90326152; PMID:2142751  
A:Accession: S10966  
A>Status: preliminary  
A:Molecule type: Protein  
A:Residues: 496-507, 'X', 509-512, 'L', 514-516, 'X', 518-519 <NAT>  
R:Plaschke-Schlutter, A.; Behrens, J.; Gherardi, E.; Birchmeier, W.  
J. Biol. Chem. 270, 830-836, 1995  
A>Title: Characterization of the scatter factor/hepatocyte growth factor gene  
A:Reference number: I48758; MUID:95122532; PMID:7822318  
A:Accession: I48758

A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-30 <RES>  
C:Cross-references: EMBL:X81630; NID:G673451; PIDN:CAA57286.1; PID:G673452  
C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor  
C:Function:  
A>Description: stimulates mitosis of hepatocytes and other cells  
A>Note: does not have proteinase activity  
C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology  
C:Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringle; pvt  
F:1-32/Domain: signal sequence #status predicted <SIG>  
F:56-495,496-728/Product: hepatocyte growth factor #status predicted <WAT>  
F:56-495/Domain: hepatocyte growth factor alpha chain #status predicted <ACH>  
F:129-207/Domain: kringle homology <KR1>  
F:129-207/Domain: kringle homology <KR2>  
F:129-207/Domain: kringle homology <KR3>  
F:306-384/Domain: kringle homology <KR4>  
F:392-470/Domain: kringle homology <KR5>  
F:496-728/Domain: hepatocyte growth factor beta chain #status predicted <BCH>  
F:496-728/Domain: hepatocyte growth factor <TRY>  
F:496-719/Domain: trypsin homology <TRY>  
F:323/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted  
F:295,403,569,656/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:489-607/Disulfide bonds: #status predicted

Query Match 17.5%; Score 139; DB 1; Length 728;  
Best Local Similarity 28.9%; Pred. No. 5.2e-05;  
Matches 43; Conservative 11; Mismatches 57; Indels 38; Gaps 11;

QY 5 HQV-PSNDC--LNGGTCVSNKYFNIHWC-----NCPKFGGQHCEIDSK 48  
Db HLTPTNFKCKDLRENYC-RNPDGASPCWFTTDNIRVGVCQIPK-----CDVSSGQ 390  
QY 49 TCYEGNGHYRKGASTDTMGRCPLPWSATVLAQTYHAH---RSDALQGLGKHYCRNP 105  
Db DCYRGNGKYNMGNLSKTRSLGTCMWKDN---MEDLHRHFWEPDASKL---NKNYCRNP 444  
QY 106 -DNRRPWCYVQGLKPLV--QECMVHDC 131  
Db 445 DDDAHGFWCYTG---NPLIPWYCPISRC 470

RESULT 48  
JC5061  
macrophage-stimulating protein 1 precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 31-Jan-1997 #sequence\_revision 31-Jan-1997 #text\_change 16-Jun-2000  
C:Accession: JC5061  
R:Ohshiro, K.; Iwama, A.; Mateuno, K.; Ezaki, T.; Sakamoto, O.; Hamaguchi, I.; Takasu, N  
Biochem. Biophys. Res. Commun. 227, 273-280, 1996  
A>Title: Molecular cloning of rat macrophage-stimulating protein and its involvement in  
A:Reference number: JC5061; MUID:97011126; PMID:8958136  
A:Accession: JC5061  
A:Molecule type: mRNA  
A:Residues: 1-716 <OHS>  
C:Cross-references: EMBL:X95096; NID:G1669718; PIDN:CAA64473.1; PID:G1669719  
C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor  
C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology  
C:Keywords: duplication; glycoprotein; growth factor; kringle  
F:1-31/Domain: signal sequence #status predicted <SIG>  
F:32-488,489-716/Product: macrophage-stimulating protein 1 #status predicted <WAT>  
F:32-488/Domain: macrophage-stimulating protein 1 alpha chain #status predicted <ACH>  
F:110-186/Domain: kringle homology <KR11>  
F:191-268/Domain: kringle homology <KR12>  
F:292-370/Domain: kringle homology <KR13>  
F:379-457/Domain: kringle homology <KR14>  
F:489-716/Domain: macrophage-stimulating protein 1 beta chain #status predicted <BCH>  
F:489-709/Domain: trypsin homology <TRY>  
F:72,305,620/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 17.4%; Score 138; DB 1; Length 716;  
Best Local Similarity 36.6%; Pred. No. 6.3e-05;  
Matches 34; Conservative 6; Mismatches 37; Indels 16; Gaps 5;

QY 30 WC---NCPKFGGQHCEIDSKTCYEG-----NGHYRKGASTDTMGRCPLPWSATVLAQ 81

Db 168 WCYTTNRSVRF--QSCGI---KSCREAVCVWNGEDYRGEVDVTESGRCQRWDLQPHS 222  
QY 82 QTYHAHRSDALQGLGKHYCRNPNDRRPWCY 114  
Db 223 HFFHPEKFPDKAL---KONYCRNPDASRPWCY 252

RESULT 49  
T18518  
apolipoprotein(a) - western European hedgehog (fragment)  
C:Species: Erinaceus europaeus (western European hedgehog)  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 07-Dec-1999  
C:Accession: T18518  
R:Lawn, R.M.; Boonmark, N.W.; Schwartz, K.; Lindahl, G.B.; Wade, D.P.; Byrne, J.; FC  
J. Biol. Chem. 270, 24004-24009, 1995  
A>Title: The recurring evolution of Lp(a): Insights from cloning of hedgehog a lipoprotein  
C:Reference number: 146259; MUID:96025778; PMID:7592597  
A:Accession: T18518  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-2869 <LAW>  
A:Cross-references: EMBL:U33170; NID:G1046358; PID:G1046359; PIDN:AAC48522.1  
A:Experimental source: liver  
C:Comment: The lipoprotein Lp(a), a major inherited risk factor for atherosclerosis, is a ent apolipoprotein(a).

Query Match 17.2%; Score 136.5; DB 2; Length 2869;  
Best Local Similarity 30.8%; Pred. No. 0.00028;  
Matches 32; Conservative 8; Mismatches 47; Indels 17; Gaps 4

QY 19 CVSNKYFNIHWCNCPKFGGQHCEIDSKT-----CYEGNGHYRKGASTDTMGSP 70  
Db 2558 CYTNSAMRWYCSIPA-----CESPTPTTEHLVPEQCLEGNGENYCGNMAITVSGGP 261  
QY 71 CLPWSATVLAQTYHAHRSDALQGLGKHYCRNPNDRRPWCY 114  
Db 2612 CQWRKQTPHREYTPENYPSKNL-FG--NYCENPGEIAPWCY 2652

RESULT 50  
S33879  
plasmin precursor - lamprey (fragments)  
N:Alternate names: plasminogen  
C:Species: Petromyzontidae gen. sp. (lamprey)  
C:Date: 02-Dec-1993 #sequence\_revision 01-Sep-1995 #text\_change 07-Nov-1997  
C:Accession: S33879  
R:Affolter, M.; Schaller, J.; Rickli, E.E.  
Protein Seq. Data Anal. 5, 207-211, 1993  
A>Title: Isolation, characterization and partial amino acid sequence of lampre  
A:Reference number: S33879  
A:Accession: S33879  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-15;16-34;35-44;45-59;60-76;77-111;111-138;139-158;159-178;179-211-217-223  
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor  
F:81-146/Domain: kringle homology <KR3>

Query Match 17.2%; Score 136; DB 2; Length 336;  
Best Local Similarity 30.7%; Pred. No. 5e-05;  
Matches 31; Conservative 7; Mismatches 41; Indels 22; Gaps 3

QY 34 PKFEGGQHCEIDSKTCYEGNGHYRKGASTDTMGRCPLPWSATVLAQTYHAHRSDALQ 93  
Db 71 PQSPAG-----LTTACVKGTGEGYRGTAAITVSGKACQAWASQT-----PGDVYS 115  
QY 94 LGLGKHYCRNPNDRRPWCYVQGLKPLVQECMVHDCADG 134  
Db 116 CQGLVSNYCRNPDEGKLPWCYT-----TEYCNVPSCGTGG 149

Search completed: May 25, 2004, 14:58:29  
Job time : 8.0894 secs

[Faint, illegible text spanning the top half of the page, possibly a header or introductory paragraph.]

---

[Faint, illegible text spanning the bottom half of the page, possibly a body paragraph or footer.]

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2004, 14:43:40 ; Search time 5.20596 Seconds  
(without alignments)  
1350.274 Million cell updates/sec

Title: US-09-880-503-4

Perfect score: 793

Sequence: 1 SNELHQVPSNCCLNGGTCV.....QVGLKPLVQCMVHDCADGX 135

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	793	100.0	431	1 UROK_HUMAN	P00749 homo sapien
2	713.5	90.0	433	1 UROK_PAPCY	P16227 papio cynoc
3	608	76.7	442	1 UROK_FIG	P04185 sus scrofa
4	579	73.0	433	1 UROK_BOVIN	Q05589 bos taurus
5	571	72.0	432	1 UROK_RAT	P29598 rattus norv
6	548	69.1	433	1 UROK_MOUSE	P06869 mus musculus
7	335.5	42.3	477	1 URT2_DESRO	P15638 desmodus ro
8	334.5	42.2	434	1 UROK_CHICK	P15120 gallus gall
9	328.5	41.4	431	1 URTB_DESRO	P08121 desmodus ro
10	328.5	41.4	562	1 TPA_HUMAN	P00750 homo sapien
11	321.5	40.5	559	1 TPA_RAT	P19637 rattus norv
12	315.5	39.8	559	1 TPA_MOUSE	P11214 mus musculus
13	310.5	39.2	477	1 URT1_DESRO	P98119 desmodus ro
14	300.5	37.9	566	1 TPA_BOVIN	Q28198 bos taurus
15	260.5	32.8	655	1 HGPA_HUMAN	Q04756 homo sapien
16	257	32.4	603	1 FA12_CAVPO	Q04962 cavia porce
17	250.5	31.6	653	1 HGPA_MOUSE	Q91098 mus musculus
18	233	29.4	615	1 FA12_HUMAN	P00748 homo sapien
19	223	28.1	394	1 URTG_DESRO	P49150 desmodus ro
20	217	27.4	593	1 FA12_BOVIN	P98140 bos taurus
21	169	21.3	685	1 ROR1_DROME	Q24488 drosophila
22	162.5	20.5	4548	1 APOA_HUMAN	P08519 homo sapien
23	161	20.3	473	1 KRMI_RAT	Q92484 rattus norv
24	161	20.3	1420	1 APOA_MACMU	P14417 macaca mula
25	159	20.1	452	1 KRMI_XENLA	Q90990 xenopus lae
26	159	20.1	473	1 KRMI_MOUSE	Q99n43 mus musculus
27	158.5	20.0	937	1 ROR1_HUMAN	Q01973 homo sapien
28	158.5	20.0	937	1 ROR1_MOUSE	Q92139 mus musculus
29	156	19.7	475	1 KRMI_HUMAN	Q96m48 homo sapien
30	156	19.7	812	1 PLMN_BOVIN	P06868 bos taurus
31	155	19.5	790	1 PLMN_FIG	P08867 sus scrofa
32	154	19.4	810	1 PLMN_HUMAN	P00747 homo sapien
33	153.5	19.4	810	1 PLMN_ERIEU	Q25485 erinaceus e

#### ALIGNMENTS

##### RESULT 1

ID	UROK_HUMAN	STANDARD;	PRT;	431 AA.
AC	P00749; Q15844; Q16618; Q969W6;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	20-MAR-1987 (Rel. 04, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)			
DE	(U-plasminogen activator).			
GN	PLAU.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OC	NCBI_TaxID=9606;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=85215647; PubMed=2987867;			
RA	Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Blasi F.;			
RT	"The human urokinase-plasminogen activator gene and its promoter.";			
RL	Nucleic Acids Res. 13:2759-2771(1985).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Holmes W.E., Pennica D., Blaber M., Rey M.W., Guenzler W.A.,			
RA	Steffens G.J., Heyneker H.L.;			
RT	"Cloning and expression of the gene for pro-urokinase in Escherichia coli.";			

P12545 macaca mula  
Q01177 rattus norv  
Q8ncw0 homo sapien  
Q01974 homo sapien  
Q92138 mus musculus  
P20318 mus musculus  
Q8k1e7 mus musculus  
P26327 homo sapien  
P14210 homo sapien  
P80009 canis famil  
P26928 mus musculus  
Q08048 mus musculus  
P17945 rattus norv  
P33574 petromyzon  
P19221 mus musculus  
P00734 homo sapien  
Q9v6k3 drosophila  
P81286 ovis aries  
P21783 xenopus lae  
P07335 bos taurus  
P18292 rattus norv  
Q9Y219 homo sapien  
Q04721 homo sapien  
Q90Y57 brachydanio  
P07898 gallus gall  
P80010 equus cabal  
Q90Y54 brachydanio  
O14594 homo sapien  
Q35516 mus musculus  
P82279 homo sapien  
Q9J171 mus musculus  
P46531 homo sapien  
P07207 drosophila  
Q9Y5W5 homo sapien  
Q9N1J7 homo sapien  
P97607 rattus norv  
Q9QW30 rattus norv  
Q07008 rattus norv  
Q9N9Y6 homo sapien  
P10040 drosophila  
Q99466 homo sapien  
Q01705 mus musculus



RL Biotechnology 3:923-929(1985).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=86056954; PubMed=2415429;

RA Nagai M., Hiramatsu R., Kaneda T., Hayasuke N., Arimura H.,

RA Nishida M., Suyama T.;

RT "Molecular cloning of cDNA coding for human preprourokinase.";

RL Gene 36:183-188(1985).

RN [4]

RP SEQUENCE FROM N.A.

RX MEDLINE=85203359; PubMed=3888571;

RA Jacobs P., Cravador A., Loriau R., Brockly P., Colau B., Chuchana P.,

RA van Elsen A., Herzog A., Bollen A.;

RT "Molecular cloning, sequencing, and expression in *Escherichia coli* of

RT human preprourokinase cDNA.";

RL DNA 4:139-146(1985).

RN [5]

RP SEQUENCE FROM N.A.

RX MEDLINE=89127526; PubMed=2536903;

RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,

RA Nickerson D.A.;

RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

RN [6]

RP SEQUENCE FROM N.A.

RC TISSUE=Lung;

RX MEDLINE=22368257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko M., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaney S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,

RA Villalón D.K., Murley D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.J., Skalska U., Smallus D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length

RT human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [7]

RP SEQUENCE OF 66-431 FROM N.A.

RX MEDLINE=84272706; PubMed=6589620;

RA Verde P., Stoppelli M.P., Galeffi P., di Nocera P., Biasi F.;

RT "Identification and primary sequence of an unspliced human urokinase

RT poly(A)+ RNA.";

RL Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).

RN [8]

RP SEQUENCE OF 21-177

RX MEDLINE=83055084; PubMed=6754569;

RA Gunzler W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E.,

RA Flohe L.;

RT "The primary structure of high molecular mass urokinase from human

RT urine. The complete amino acid sequence of the A chain.";

RL Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).

RN [9]

RP SEQUENCE OF 156-176 AND 179-224.

RX MEDLINE=83003608; PubMed=6749491;

RA Schaller J., Nick H., Rickli E.B., Gillesen D., Lergier W.,

RA Studer R.O.;

RT "Human low-molecular-weight urinary urokinase. Partial

RT characterization and preliminary sequence data of the two polypeptide

RT chains.";

RL Eur. J. Biochem. 125:251-257(1982).

RN [10]

RP SEQUENCE OF 158-410.

RX MEDLINE=83055099; PubMed=6754572;

RA Steffens G.J., Gunzler W.A., Otting F., Frankus E., Flohe L.;

RT "The complete amino acid sequence of low molecular mass urokinase

RT from human urine.";

RL Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).

RN [11]

RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).

RX MEDLINE=96000858; PubMed=8591045;

RA Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,

RA Dobson C.M., Stuart D.I., Jones E.Y.;

RT "The crystal structure of the catalytic domain of human

RT urokinase-type plasminogen activator.";

RL Structure 3:681-691(1995).

RN [12]

RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.

RX MEDLINE=20266327; PubMed=10805774;

RA Sperl S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G.,

RA Bode W., Magdolen V., Huber R., Moroder L.;

RT "(4-aminomethyl)phenylguanidine derivatives as nonpeptidic highly

RT selective inhibitors of human urokinase.";

RL Proc. Natl. Acad. Sci. U.S.A. 97:5113-5116(2000).

RN [13]

RP STRUCTURE BY NMR.

RX MEDLINE=89127526; PubMed=2536903;

RA Oswald R.E., Bogusky M.J., Bamberger M., Smith R.A.G., Dobson C.M.;

RT "Dynamics of the multidomain fibrinolytic protein urokinase from two-

RT dimensional NMR.";

RL Nature 337:579-582(1989).

RN [14]

RP STRUCTURE BY NMR OF 67-155.

RX MEDLINE=93003110; PubMed=1327118;

RA Li X., Smith R.A.G., Dobson C.M.;

RT "Sequential 1H NMR assignments and secondary structure of the kringie

RT domain from urokinase.";

RL Biochemistry 31:9562-9571(1992).

RN [15]

RP STRUCTURE BY NMR OF 67-155.

RX MEDLINE=94149701; PubMed=8107091;

RA Li X., Bokman A.M., Llinas M., Smith R.A.G., Dobson C.M.;

RT "Solution structure of the kringie domain from urokinase-type

RT plasminogen activator.";

RL J. Mol. Biol. 235:1548-1559(1994).

RN [16]

RP VARIANT LEU-141.

RX MEDLINE=96186279; PubMed=8652631;

RA Yoshimoto M., Ushiyama Y., Sakai M., Tamaki S., Hara H., Takahashi K.,

RA Sawasaki Y., Hanada K.;

RT "Characterization of single chain urokinase-type plasminogen

RT activator with a novel amino-acid substitution in the kringie

RT structure.";

RL Biochim. Biophys. Acta 1293:83-89(1996).

RN [17]

RP VARIANT LEU-141.

RX MEDLINE=97218551; PubMed=9065988;

RA Conne B., Berczy M., Belin D.;

RT "Detection of polymorphisms in the human urokinase-type plasminogen

RT activator gene.";

RL Thromb. Haemost. 77:434-435(1997).

RN [18]

RP ERRATUM.

RA Conne B., Berczy M., Belin D.;

RA Thromb. Haemost. 78:973-973(1997).

RN [19]

RP VARIANT LEU-141.

RX MEDLINE=97337920; PubMed=9194591;

RA Turkmen B., Schmitt M., Schmalfeldt B., Trommler P., Hell W.,

RA Creutzburg S., Graeff H., Magdolen V.;

RT "Mutational analysis of the genes encoding urokinase-type plasminogen

RT activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer.";

RL Electrophoresis 18:686-689(1997).

CC -!- FUNCTION: Potent plasminogen activator and is clinically used for

CC therapy of thrombolytic disorders.

CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in

CC plasminogen to form plasmin.

CC -!- SUBUNIT: Found in high and low molecular mass forms. Each consists

of two chains, A and B. The high molecular mass form contains a long chain A. Cleavage occurs after residue 155 in the low molecular mass form to yield a short A1 chain.

PHARMACEUTICAL: Available under the name Abbokinase (Abbott). Used in Pulmonary Embolism (PE) to initiate fibrinolysis.

SIMILARITY: Belongs to peptidase family S1.

SIMILARITY: Contains 1 EGF-like domain.

SIMILARITY: Contains 1 kringle domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; X02419; CAA26268.1; -  
EMBL; M15476; AAA61253.1; -  
EMBL; D00244; BAA00175.1; -  
EMBL; D11143; BAA01819.1; -  
EMBL; X02760; CAA26535.1; -  
EMBL; AF377330; AAK53822.1; -  
EMBL; BC013575; AAH13575.1; -  
EMBL; K03226; AAC97138.1; -  
EMBL; K02286; AAA61252.1; -  
EMBL; A21571; CAA01559.1; -  
EMBL; A18397; CAA01390.1; -  
PIR; A00931; URGU.  
PDB; 1KDUJ; 31-OCT-93.

Query Match 100.0%; Score 793; DB 1; Length 431;  
Best Local Similarity 100.0%; Pred. No. 1.5e-70;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCCLNGGTCVSNKYFSNIHWNCNPKKFGQGCCEIDKSKTCYEGNGHFYRG 60  
DB 21 SNELHQVPSNCCLNGGTCVSNKYFSNIHWNCNPKKFGQGCCEIDKSKTCYEGNGHFYRG 80  
QY 61 KASTDTMGPCPLPNSATVLOQTYHAHRSALQLGLGKHNCNPNDRRRPCVQVGLK 120  
DB 81 KASTDTMGPCPLPNSATVLOQTYHAHRSALQLGLGKHNCNPNDRRRPCVQVGLK 140  
QY 121 PLVQECMVHDCADGK 135  
DB 141 PLVQECMVHDCADGK 155

RESULT 2  
UROK\_PAPCY STANDARD; PRT; 433 AA.  
AC P16227;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)  
DE (U-plasminogen activator).  
GN FLAU.  
OS Papio cynocephalus (Yellow baboon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;  
OC Cercopitheidae; Papio.  
OX NCBI\_TaxID=9556;  
RN [1]\_TaxID=9556;  
RP SEQUENCE FROM N.A.  
RC TISSUE=Thoracic aorta;  
RX MEDLINE=90287734; PubMed=2113276;  
RA Au Y.P.T., Wang T.W., Clowes A.W.;  
RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminogen activator";  
RT Nucleic Acids Res. 18:3411-3411(1990).  
RL -/- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-val bond in plasminogen to form plasmin.

-/- SUBUNIT: Found in high and low molecular mass forms. Each consists of two chains, A and B. The high molecular mass form contains a long chain A. Cleavage occurs after residue 155 in the low molecular mass form to yield a short A1 chain (By similarity).

-/- SIMILARITY: Belongs to peptidase family S1.

-/- SIMILARITY: Contains 1 EGF-like domain.

-/- SIMILARITY: Contains 1 kringle domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; X51935; CAA36200.1; -  
PIR; S14687; UKBAY.  
HSP; P00749; ILMW.  
MEROPS; S01.231; -  
InterPro; IPR009003; Cys\_Ser\_trypsin.  
InterPro; IPR006209; EGF\_like.  
InterPro; IPR006210; IEGF.  
InterPro; IPR000001; Kringle.  
InterPro; IPR008293; Pept\_S1A\_UPA.  
InterPro; IPR001254; Peptidase\_S1.  
InterPro; IPR001314; Peptidase\_S1A.  
Pfam; PF00051; kringle; 1.  
Pfam; PF00089; trypsin; 1.  
PIRSF; PIRSF001144; UPK\_Plaem\_act; 1.  
PRINTS; PR00722; CHYMOTRYPSIN.  
PRINTS; PR00018; KRINGLE.  
ProDom; PD000395; Kringle; 1.  
SMART; SM00181; EGF; 1.  
SMART; SM00130; KR; 1.  
SMART; SM00020; Tryp\_Spc; 1.  
PROSITE; PS00022; EGF\_1; 1.  
PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
PROSITE; PS00026; EGF\_3; 1.  
PROSITE; PS00021; KRINGLE\_1; 1.  
PROSITE; PS00070; KRINGLE\_2; 1.  
PROSITE; PS02040; TRYPSIN\_DOM; 1.  
PROSITE; PS00134; TRYPSIN\_HIS; 1.  
PROSITE; PS00135; TRYPSIN\_SER; 1.  
PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein; Kringle; EGF-like domain; Zymogen; Signal.  
FT SIGNAL 1 20  
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.  
FT CHAIN 21 176 CHAIN A (BY SIMILARITY).  
FT CHAIN 155 176 SHORT A CHAIN (A1) (BY SIMILARITY).  
FT CHAIN 178 433 CHAIN B (BY SIMILARITY).  
FT DOMAIN 26 62 EGF-LIKE.  
FT DOMAIN 69 150 KRINGLE.  
FT DOMAIN 151 177 CONNECTING PEPTIDE.  
FT DOMAIN 178 433 SERINE PROTEASE.  
FT DISULFID 30 38 BY SIMILARITY.  
FT DISULFID 32 50 BY SIMILARITY.  
FT DISULFID 52 61 BY SIMILARITY.  
FT DISULFID 167 298 INTERCHAIN (BY SIMILARITY).  
FT DISULFID 208 224 BY SIMILARITY.  
FT DISULFID 216 287 BY SIMILARITY.  
FT DISULFID 315 384 BY SIMILARITY.  
FT DISULFID 347 363 BY SIMILARITY.  
FT DISULFID 374 402 BY SIMILARITY.  
FT ACT\_SITE 223 223 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 274 274 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 378 378 CHARGE RELAY SYSTEM.  
FT CARBOHYD 324 324 N-LINKED (GLCNAC...) (BY SIMILARITY).  
SQ SEQUENCE 433 AA; 48595 MW; 816D22DFDDC8792 CRC64;

Query Match 90.0%; Score 713.5; DB 1; Length 433;  
Best Local Similarity 91.9%; Pred. No. 9.2e-63;  
Matches 124; Conservative 4; Mismatches 6; Indels 1; Gaps 1

QY 1 SNELHGVPSNCDLNGGTCVSNKYFSNIHWNCBPKFGGHCEDKSKTCYEGNGHFYRG 60  
 DB 21 SREL-QVPSDCGCLNGGTCMKNKFSSIHWNCBPKFGGHCEDKSKTCYEGNGHFYRG 79  
 QY 61 KASDTDMGRPCLPWNSATVLTQTYHAHRSALQGLGKHNYCRPNDRPWCYVQVGLK 120  
 DB 80 KASDTDMGRSCLANWSATVLTQTYHAHRSALQGLGKHNYCRPNDRPWCYVQVGLK 139  
 QY 121 PLVQECMVHDCADGK 135  
 DB 140 QRVCQCMVHNCADGK 154

RESULT 3  
 UROK\_PIG STANDARD; PRT; 442 AA.  
 AC F04185;  
 DT 20-MAR-1987 (Rel. 04, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)  
 DE (U-plasminogen activator).  
 GN PLAU.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=85987954; PubMed=6096832;  
 RA Nagamine Y., Pearson D., Alcus M.S., Reich E.;  
 RT "cDNA and gene nucleotide sequence of porcine plasminogen activator.";  
 RL Nucleic Acids Res. 12:9525-9541 (1984).  
 RN [2]  
 RP REVISION TO 241.  
 RA Nagamine Y.  
 RL Submitted (DEC-1986) to the PIR data bank.  
 CC -|- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
 CC plasminogen to form plasmin.  
 CC -|- SIMILARITY: Belongs to peptidase family S1.  
 CC -|- SIMILARITY: Contains 1 EGF-like domain.  
 CC -|- SIMILARITY: Contains 1 kringle domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X01648; CAA25806.1; -;  
 CC EMBL; X02724; CAA26511.1; -;  
 CC PIR; A00932; URPG.  
 CC HSP; P00749; IKDU.  
 CC MROPS; S01.231; -;  
 CC InterPro; IPR009003; Cys Ser trypsin.  
 CC InterPro; IPR006203; EGF-like.  
 CC InterPro; IPR000001; Kringle.  
 CC InterPro; IPR008293; Pept\_S1A\_uPA.  
 CC InterPro; IPR001254; Peptidase\_S1.  
 CC InterPro; IPR001314; Peptidase\_S1A.  
 CC Pfam; PF00051; kringle; 1.  
 CC Pfam; PF00089; trypsin; 1.  
 CC PIRSF; PIRSF001144; Urk plasmin act; 1.  
 CC PRINTS; PR00722; CHYMOTRYPsin.  
 CC PRINTS; PR00018; KRINGLE.  
 CC ProDom; PD000395; Kringle; 1.  
 CC SMART; SMC0130; KR; 1.  
 CC SMART; SMC0020; Tryp\_Spc; 1.  
 CC PROSITE; PS00022; EGF\_1; 1.

DR PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
 DR PROSITE; PS00026; EGF\_3; 1.  
 DR PROSITE; PS00021; KRINGLE\_1; 1.  
 DR PROSITE; PS00070; KRINGLE\_2; 1.  
 DR PROSITE; PS00240; TRYPsin\_DOM; 1.  
 DR PROSITE; PS00134; TRYPsin\_HIS; 1.  
 DR PROSITE; PS00135; TRYPsin\_SER; 1.  
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
 KW Kringle; EGF-like domain; Zymogen; Signal.  
 FT SIGNAL 1 20  
 FT CHAIN 21 442 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.  
 FT CHAIN 21 188 CHAIN A (BY SIMILARITY).  
 FT CHAIN 190 442 CHAIN B (BY SIMILARITY).  
 FT DOMAIN 29 65 EGF-LIKE.  
 FT DOMAIN 72 153 KRINGLE.  
 FT DOMAIN 154 189 CONNECTING PEPTIDE.  
 FT DOMAIN 190 442 SERINE PROTEASE.  
 FT CARBOHYD 152 152 N-LINKED (GLCNAC...).  
 FT DISULFID 33 41 BY SIMILARITY.  
 FT DISULFID 35 53 BY SIMILARITY.  
 FT DISULFID 55 64 BY SIMILARITY.  
 FT DISULFID 179 310 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 220 236 BY SIMILARITY.  
 FT DISULFID 228 299 BY SIMILARITY.  
 FT DISULFID 324 393 BY SIMILARITY.  
 FT DISULFID 356 372 BY SIMILARITY.  
 FT DISULFID 383 411 BY SIMILARITY.  
 FT ACT\_SITE 235 286 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 286 286 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 287 387 CHARGE RELAY SYSTEM.  
 FT CONFLICT 241 241 Q -> H (IN REF. 1; CAA25806).  
 FT CONFLICT 242 242 Q -> H (IN REF. 1; CAA26511).  
 FT CONFLICT 288 288 A -> GS (IN REF. 1; CAA25806).  
 SQ SEQUENCE 442 AA; 49116 MW; EE32FCBF501321EE CRC64;  
 Query Match 76.7%; Score 608; DB 1; Length 442;  
 Best local Similarity 78.1%; Pred. No. 2.1e-52;  
 Matches 107; Conservative 11; Mismatches 17; Indels 2; Gaps 1

QY 1 SNELHGV--PSNCDLNGGTCVSNKYFSNIHWNCBPKFGGHCEDKSKTCYEGNGHFY 58  
 DB 21 SRELHVESGASNGCGLNGKCVSYKVFNSIQRCSPKFKQGEHCEDTQTCTFEGNHYSY 80  
 QY 59 RGKASDTDMGRPCLPWNSATVLTQTYHAHRSALQGLGKHNYCRPNDRPWCYVQV 118  
 DB 81 RGRANTNTGGRPCLPWNSATVLTNTYHAHRPDALQGLGKHNYCRPNDRPWCYVQV 140  
 QY 119 LKQLVQECMVHDCADGK 135  
 DB 141 LKQLVQECMVHNCADGK 157

RESULT 4  
 UROK\_BOVIN STANDARD; PRT; 433 AA.  
 AC Q05589; Q28209;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)  
 DE (U-plasminogen activator).  
 GN PLAU.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Aortic endothelium;  
 RX MEDLINE=93216119; PubMed=8385052;  
 RA Kraetzschmar J., Haendler B., Kojima S., Rifkin D.B.,  
 RA Schleuning W.-D.;

FT ACT\_SITE 378 378 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT CONFLICT 189 189 A -> T (IN REF. 2).  
 FT SEQUENCE 433 AA; 48730 MW; 4DE1B8D4D47027A CRC64;  
 SQ

Query Match 73.0%; Score 579; DB 1; Length 433;  
 Best Local Similarity 73.0%; Pred. No. 1.4e-49;  
 Matches 100; Conservative 13; Mismatches 22; Indels 2; Gaps 1

QY 1 SNEHQV--PSNCDCLNGTCVSNKYFSNIHWCNCPKFGGCHCEIDKSKTCYEGNGHY 58  
 Db 21 SNEVHKESGESNCGCLNGKCVYFYFSNIQRCSCPKFGGCHCEIDKSKTCYQNGHSHY 80  
 QY 59 RKGASTDTMTGRPCLPWNSATVLOQTYHAHRSALQGLGKHNCRPNRRPWCYQVQ 118  
 Db 81 RKGANRDLGRPCLAWSPTVLLKMYHAHRSALQGLGKHNCRPNRRPWCYQVQ 140  
 QY 119 LKPLVQECMWHDCADGK 135  
 Db 141 LKQFVQFCMVQDCSVGK 157

RESULT 5  
 UROK\_RAT STANDARD; PRT; 432 AA.  
 ID UROK\_RAT  
 AC P29598;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)  
 DE (U-plasminogen activator).  
 GN PLAU.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus;  
 CX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Fischer 344;  
 RX MEDLINE=92233409; PubMed=1568219;  
 RA Henderson B.R.; Tansey W.P.; Phillips S.M.; Ramshaw I.A.;  
 RA Kefford R.F.;  
 RT "Transcriptional and posttranscriptional activation of urokinase  
 RT plasminogen activator gene expression in metastatic tumor cells";  
 RL Cancer Res. 52:2489-2496(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Rabbani S.A.;  
 RA Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg|-Val bond in  
 CC plasminogen to form plasmin.  
 CC -1- SUBUNIT: Found in high and low molecular mass forms. Each consists  
 CC of two chains, A and B. The high molecular mass form contains a  
 CC long chain A. Cleavage occurs after residue 156 in the low  
 CC molecular mass form to yield a short A1 chain (By similarity).  
 CC -1- SIMILARITY: Belongs to peptidase family S1.  
 CC -1- SIMILARITY: Contains 1 EGF-like domain.  
 CC -1- SIMILARITY: Contains 1 kringle domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboratio  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on it  
 CC use by non-profit institutions as long as its content is in no wa  
 CC modified and this statement is not removed. Usage by and for commercia  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X63434; CAA45028.1; -;  
 CC EMBL; X65651; CAA46601.1; -;  
 CC PIR; S24604; S18932.  
 CC HSSP; P00749; 1KDU.  
 CC DR MEROPS; S01.231; -;  
 CC InterPro; IPR009003; Cys\_Ser\_trypsin.

"Bovine urokinase-type plasminogen activator and its receptor:  
 cloning and induction by retinoic acid.";  
 RL Gene 125:177-183(1993).  
 RN [2]  
 RP SEQUENCE OF 12-433 FROM N.A.  
 RC TISSUE=Kidney;  
 RA Rayn P.; Berglund L.; Petersen T.E.;  
 RT "Cloning and characterization of the bovine plasminogen activators uPA  
 RT and tPA";  
 RL Int. Dairy J. 5:605-617(1995).  
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg|-Val bond in  
 CC plasminogen to form plasmin.  
 CC -1- INDUCTION: By retinoic acid.  
 CC -1- SIMILARITY: Belongs to peptidase family S1.  
 CC -1- SIMILARITY: Contains 1 EGF-like domain.  
 CC -1- SIMILARITY: Contains 1 kringle domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboratio  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; L03546; AAA51419.1; -;  
 CC EMBL; X85801; CAA59796.1; -;  
 CC PIR; JN0560; JN0560.  
 CC HSSP; P00749; 1LMW.  
 CC DR MEROPS; S01.231; -;  
 CC InterPro; IPR009003; Cys\_Ser\_trypsin.  
 CC InterPro; IPR006209; EGF\_like.  
 CC InterPro; IPR000001; Kringle.  
 CC InterPro; IPR008293; Pept\_S1A\_uPA.  
 CC InterPro; IPR001254; Peptidase\_S1.  
 CC InterPro; IPR001314; Peptidase\_S1A.  
 CC Pfam; PF00051; kringle; 1.  
 CC Pfam; PF00889; trypsin; 1.  
 CC PIRSF; PIRSF001144; Uro\_kinase\_act; 1.  
 CC PRINTS; PR00722; CHYNOTRYPsin.  
 CC PRINTS; PR00018; KRINGLE.  
 CC ProDom; PD000395; Kringle; 1.  
 CC SMART; SM00130; KR; 1.  
 CC SMART; SM00020; TRYD\_SPC; 1.  
 CC PROSITE; PS00022; EGF\_1; 1.  
 CC PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
 CC PROSITE; PS00026; EGF\_3; 1.  
 CC PROSITE; PS00021; KRINGLE\_1; 1.  
 CC PROSITE; PS00070; KRINGLE\_2; 1.  
 CC PROSITE; PS0240; TRYPSIN\_DOM; 1.  
 CC PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 CC PROSITE; PS00135; TRYPSIN\_SER; 1.  
 CC Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
 CC Kringle; EGF-like domain; Signal; Zymogen.  
 FT SIGNAL 1 20  
 FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.  
 FT CHAIN 21 179 CHAIN A (BY SIMILARITY).  
 FT CHAIN 181 433 CHAIN B (BY SIMILARITY).  
 FT DOMAIN 29 65 EGF-LIKE.  
 FT DOMAIN 72 153 KRINGLE.  
 FT DOMAIN 154 180 CONNECTING PEPTIDE.  
 FT DOMAIN 181 433 SERINE PROTEASE.  
 FT DISULFID 33 41 BY SIMILARITY.  
 FT DISULFID 35 53 BY SIMILARITY.  
 FT DISULFID 55 64 BY SIMILARITY.  
 FT DISULFID 170 301 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 211 227 BY SIMILARITY.  
 FT DISULFID 219 290 BY SIMILARITY.  
 FT DISULFID 315 384 BY SIMILARITY.  
 FT DISULFID 347 363 BY SIMILARITY.  
 FT DISULFID 374 402 BY SIMILARITY.  
 FT ACT\_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 227 277 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 277

01-JAN-1988 (Rel. 06, Last sequence update)  
28-FEB-2003 (Rel. 41, Last annotation update)  
Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)  
(U-plasminogen activator).  
PLAU.  
GN  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RN  
SEQUENCE FROM N.A.  
RP MEDLINE=85179474; PubMed=298383;  
RX MEDLINE=85179474; PubMed=298383;  
RX Belin D., Vassalli J.-D., Conbepine C., Godeau F., Nagamine Y.,  
RA Reich E., Kocher H.P., Duvoisin R.M.;  
RA "Cloning, nucleotide sequencing and expression of cDNAs encoding  
RT mouse urokinase-type plasminogen activator.";  
RT Eur. J. Biochem. 148:225-232(1985).  
RL [2]  
RN  
SEQUENCE FROM N.A.  
RP MEDLINE=85163489; PubMed=2831940;  
RX Degen S.J.F., Heckel J.L., Reich E., Degen J.L.;  
RA "The murine urokinase-type plasminogen activator gene.";  
RT Biochemistry 26:8270-8279(1987).  
CC -|- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
CC plasminogen to form plasmin.  
CC -|- SUBUNIT: Found in high and low molecular mass forms. Each consists  
CC of two chains, A and B. The high molecular mass form contains a  
CC long chain A. Cleavage occurs after residue 156 in the low  
CC molecular mass form to yield a short A1 chain (By similarity).  
CC -|- SIMILARITY: Belongs to peptidase family S1.  
CC -|- SIMILARITY: Contains 1 EGF-like domain.  
CC -|- SIMILARITY: Contains 1 kringle domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaborative  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on it  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement>  
CC or send an email to [license@sb-sib.ch](mailto:license@sb-sib.ch)).  
CC  
CC  
CC EMBL: X02389; CAA26231.1; -  
DR EMBL: M17922; AAA40539.1; -  
DR FIR; A29420; UNKS.  
DR HSSP; P00749; IXDU.  
DR MEROPS; S01.231; -.  
DR MGD; MG1.97611; PLAU.  
DR InterPro; IPR009003; Cys Ser trypsin.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR008293; Pept\_S1A\_uPA.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00051; kringle; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PIRSF; PIRSF001144; Urk plasm act; 1.  
DR PRINTS; P00722; CHYMOTRYPSIN.  
DR PRINTS; P00018; KRINGLE.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00181; EGF; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
DR PROSITE; PS50026; EGF\_3; 1.  
DR PROSITE; PS50021; KRINGLE\_1; 1.  
DR PROSITE; PS50021; KRINGLE\_2; 1.  
DR PROSITE; PS50070; KRINGLE\_2; 1.  
DR PROSITE; PS50240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS50134; TRYPSIN\_HIS; FALSE\_NEG.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Kringle; EGF-like domain; Zymogen; signal.  
KW Plasmigen activation; Hydrolase; Serine protease; Glycoprotein;

FT	SIGNAL	1	20	POTENTIAL.
FT	CHAIN	21	433	UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT	CHAIN	21	178	CHAIN A (BY SIMILARITY).
FT	CHAIN	157	178	SHORT A CHAIN (A1).
FT	CHAIN	180	433	CHAIN B (BY SIMILARITY).
FT	DOMAIN	28	64	EGF-LIKE.
FT	DOMAIN	71	152	KRINGLE.
FT	DOMAIN	153	179	CONNECTING PEPTIDE.
FT	DOMAIN	180	433	SERINE PROTEASE.
FT	DISULFID	32	40	BY SIMILARITY.
FT	DISULFID	34	52	BY SIMILARITY.
FT	DISULFID	54	63	BY SIMILARITY.
FT	DISULFID	159	301	INTERCHAIN (BY SIMILARITY).
FT	DISULFID	211	227	BY SIMILARITY.
FT	DISULFID	219	290	BY SIMILARITY.
FT	DISULFID	315	384	BY SIMILARITY.
FT	DISULFID	347	363	BY SIMILARITY.
FT	DISULFID	374	402	BY SIMILARITY.
FT	ACT_SITE	226	226	CHARGE RELAY SYSTEM.
FT	ACT_SITE	277	277	CHARGE RELAY SYSTEM.
FT	ACT_SITE	378	378	CHARGE RELAY SYSTEM.
ST	SEQUENCE	433 AA;	48268 MW; A99C3SF6250443F9 CRC64;	

  

Query Match		69.1%;	Score 548;	DB 1;	Length 433;
Best Local Similarity		71.7%;	Pred. No. 1.6e-46;		
Matches 91;		Conservative 13;	Mismatches 23;	Indels 0;	Gaps 0;

  

QY	9	SNCDCLNGGTCVSNKYPSNIHWNCNPKFGQHCEIDKSKTCYEGHGFYRGKASTDTWG	68
DB	30	SNCGCGNGVCVSXYKFIRRCRCPRKFGHEHCIDASKTCYHGNGDSYRGKANTDTKG	89
QY	69	RPLCPWNSATVLQQTTHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVLKPLVQECNV	128
DB	90	RECLAWNAPAVLQKPYNHAHRDAISLGKHNYCRNPDKNKRPWCYVQIGLQFVQECNV	149
QY	129	HDCADGK	135
DB	150	HDCSLSK	156

  

RESULT 7

URP2_DESRO		STANDARD;	PRT;	477 AA.
ID	URT2_DESRO			
AC	F15638;			
DT	01-APR-1990 (Rel. 14, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Salivary plasminogen activator alpha 2 precursor (EC 3.4.21.68) (DSPA			
DE	alpha-2) (BAT-PA) (T-plasminogen activator).			
OS	Desmodus rotundus (Vampire bat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;			
OC	Desmodontinae; Desmodus.			
OX	NCBI_TaxID=9430;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Salivary gland;			
RX	MEDLINE=92039036; PubMed=1937019;			
RA	Kraetzschmar J., Haendler B., Langer G., Boide W., Bringmann P.,			
RA	Alagon A., Donner P., Schleuning W.D.;			
RT	"The plasminogen activator family from the salivary gland of the			
RT	vampire bat Desmodus rotundus: cloning and expression."			
RL	Gene 105:229-237(1991).			
RN	[2]			
RP	SEQUENCE FROM N.A.; AND PARTIAL SEQUENCE.			
RC	TISSUE=Salivary gland;			
RX	MEDLINE=90036867; PubMed=2509450;			
RA	Gardell S.J., Duong L.T., Diehl R.E., York J.D., Hare T.R.,			
RA	Register R.B., Jacobs J.W., Dixon R.A.F., Friedman P.A.;			
RT	"Isolation, characterization, and cDNA cloning of a vampire bat			
RT	salivary plasminogen activator."			
RN	J. Biol. Chem. 264:17947-17952(1989).			
RN	[3]			

FT ACT\_SITE 272 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT\_SITE 321 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT\_SITE 428 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT DISULFID 42 BY SIMILARITY.

FT DISULFID 70 BY SIMILARITY.

FT DISULFID 87 BY SIMILARITY.

FT DISULFID 92 BY SIMILARITY.

FT DISULFID 111 BY SIMILARITY.

FT DISULFID 128 BY SIMILARITY.

FT DISULFID 149 BY SIMILARITY.

FT DISULFID 180 BY SIMILARITY.

FT DISULFID 214 BY SIMILARITY.

FT DISULFID 257 BY SIMILARITY.

FT DISULFID 265 BY SIMILARITY.

FT DISULFID 359 BY SIMILARITY.

FT DISULFID 391 BY SIMILARITY.

FT DISULFID 424 BY SIMILARITY.

FT CARBOHYD 185 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 398 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CONFLICT 403 N -> K (IN REF. 2).

FT CONFLICT 417 Y -> H (IN REF. 2).

FT CONFLICT 435 M -> R (IN REF. 2).

FT SEQUENCE 477 AA; 53719 MW; 17486555COE5077C CRC64;

Query Match 42.3%; Score 335.5; DB 1; Length 477;

Best Local Similarity 46.3%; Pred. No. 1.2e-25;

Matches 62; Conservative 17; Mismatches 50; Indels 5; Gaps 2;

QY 3 ELHQVP- - - - - SNCDCLNGGTCVSNKYFVNIHWCNPKKFGGQHCIDSKTCYEGNGHFY 58

78 QCHTVPVKSCLRFNGTCTQWASFSDF-VQCPKGYTGKQCEVDTHATCYKDGVTY 136

QY 59 RKASDTNWRCLPNSATVQOTYHAHRSALQGLGKHNCRPNRRRPPWCYQVG 118

137 RGTWSTESGAQICWNSNLLTRTYNGRRSDAITGLGNHNYCRPNPNNSKPCWYVTKA 196

QY 119 LKPLVCEQMVHCA 132

197 SKFILEFCVPVCS 210

DB

RESULT 8

UOK CHICK STANDARD; PRT; 434 AA.

AC P15120; 1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)

DE (U-plasminogen activator).

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OX NCBI\_TaxID=9031;

RN [1]

SEQUENCE FROM N.A.

RP MEDLINE=90110185; PubMed=2295632;

RA Leslie N.D., Kessler C.A., Bell S.M., Degen J.L.;

RT "The chicken urokinase-type plasminogen activator gene.;"

RL J. Biol. Chem. 265:1339-1344(1990).

CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.

CC -1- SIMILARITY: Belongs to peptidase family S1.

CC -1- SIMILARITY: Contains 1 EGF-like domain.

CC -1- SIMILARITY: Contains 1 kringle domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

CC

CC ENBL; J05187; AAA49131.1; -

DR ENBL; J05188; AAA49130.1; -

DR PIR; A35005; A35005.

DR HSP; P00763; IDPO.

DR MEROPS; S01.231; -

DR InterPro; IPR009003; Cys Ser trypsin.

DR InterPro; IPR006209; EGF-like.

DR InterPro; IPR006210; IEGF.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR008293; Pept\_S1A\_uPA.

DR InterPro; IPR001254; Peptidase\_S1.

DR InterPro; IPR001314; Peptidase\_S1A.

DR Pfam; PF00051; Kringle; 1.

DR Pfam; PF00089; trypsin; 1.

DR PRSE; PRSF001144; Urk\_plasm\_act; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PR00018; KRINGLE.

DR ProDom; PD000395; Kringle; 1.

DR SMART; SM00181; EGF; 1.

DR SMART; SM00130; KR; 1.

DR SMART; SM00020; Tryp\_SPC; 1.

DR PROSITE; PS00022; EGF\_1; 1.

DR PROSITE; PS01186; EGF\_2; 1.

DR PROSITE; PS00026; EGF\_3; 1.

DR PROSITE; PS00021; KRINGLE\_1; 1.

DR PROSITE; PS00021; KRINGLE\_2; 1.

DR PROSITE; PS00070; TRYPsin; 1.

DR PROSITE; PS00240; TRYPsin; 1.

DR PROSITE; PS00134; TRYPsin; 1.

DR PROSITE; PS00135; TRYPsin; 1.

DR Plasminogen activation; Hydrolase, Serine protease; Glycoprotein; Kringle; EGF-like domain; Signal; Zymogen.

KW SIGNAL

KW FT CHAIN 21 434 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.

FT CHAIN 21 171 CHAIN A (BY SIMILARITY).

FT CHAIN 173 434 CHAIN B (BY SIMILARITY).

FT DOMAIN 36 72 EGF-LIKE.

FT DOMAIN 79 158 KRINGLE.

FT DOMAIN 159 172 CONNECTING PEPTIDE.

FT DOMAIN 173 434 SERINE PROTEASE.

FT DISULFID 40 48 BY SIMILARITY.

FT DISULFID 42 60 BY SIMILARITY.

FT DISULFID 62 71 INTERCHAIN (BY SIMILARITY).

FT DISULFID 162 296 BY SIMILARITY.

FT DISULFID 202 218 BY SIMILARITY.

FT DISULFID 210 285 BY SIMILARITY.

FT DISULFID 310 379 BY SIMILARITY.

FT DISULFID 342 358 BY SIMILARITY.

FT DISULFID 369 397 BY SIMILARITY.

FT ACT\_SITE 217 217 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT\_SITE 272 272 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT\_SITE 373 373 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT CARBOHYD 228 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 434 AA; 49400 MW; BD861048DD66A55 CRC64;

Query Match 42.2%; Score 334.5; DB 1; Length 434;

Best Local Similarity 54.2%; Pred. No. 1.4e-25;

Matches 58; Conservative 14; Mismatches 30; Indels 5; Gaps 2

QY 11 CDCCLNGGTCVSNKYFVNIHWCNPKKFGGQHCIDSKTCYEGNGHFYRGKASTDTMGRP 70

DB 40 CQCLNGGTCITYRFFSQKCLCEGEGGJHCIDTNSICYSNGEDYRGNAEDP-----G 95

QY 71 CLPWSATVLQ-QTYHAHRSALQGLGKHNCRPNRRRPPWCYQ 116

DB 96 CLYWDHPSVIRWGVDHADLNALQLGLGKHNCRPNNGRSRPPWCYTK 142

RESULT 9

URTE DESRO

ID URTE DESRO STANDARD; PRT; 431 AA.

AC P98121;

01-FEB-1996 (Rel. 33, Created)  
01-FEB-1996 (Rel. 33, Last sequence update)  
28-FEB-2003 (Rel. 41, Last annotation update)  
Salivary plasminogen activator beta precursor (EC 3.4.21.68) (DSPA beta).  
Desmodus rotundus (Vampire bat).  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae; Desmodontinae; Desmodus.  
NCBI\_TaxID=9430;  
[1]  
SEQUENCE FROM N.A.  
TISSUE=Salivary Gland;  
MEDLINE=92039036; PubMed=1937019;  
Alagon A., Donner P., Haendler B., Langer G., Boidol W., Bringmann P., Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P., Donner P.,  
"The plasminogen activator family from the salivary gland of the vampire bat Desmodus rotundus: cloning and expression.";  
Gene 105:229-237(1991).  
[2]  
CHARACTERIZATION.  
MEDLINE=93393059; PubMed=1309059;  
Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T., Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P., Donner P.,  
"Plasminogen activators from the saliva of Desmodus rotundus (common vampire bat): unique fibrin specificity.";  
Ann. N.Y. Acad. Sci. 667:995-1003(1992).  
-!- FUNCTION: Probably essential to support the feeding habits of this exclusively haematophagous animal. Probable potent thrombolytic agent.  
-!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.  
-!- SUBUNIT: Monomer.  
-!- SIMILARITY: Belongs to peptidase family S1.  
-!- SIMILARITY: Contains 1 EGF-like domain.  
-!- SIMILARITY: Contains 1 kringle domain.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; M63889; AAA31594.1; -  
PIR; J050599; J050599.  
HSP; P98119; IAS1.  
MEROPS; S01.239; -  
InterPro; IPR009003; Cys Ser trypsin.  
InterPro; IPR006209; EGF-like.  
InterPro; IPR006210; IEGF.  
InterPro; IPR000001; Kringle.  
InterPro; IPR001254; Peptidase S1.  
InterPro; IPR001314; Peptidase\_S1A.  
Pfam; PF00008; EGF; 1.  
Pfam; PF00051; Kringle; 1.  
Pfam; PF00089; trypsin; 1.  
PRINTS; PR00722; CHYMOTRYPSIN.  
PRINTS; PR00016; KRINGLE.  
ProDom; PD000395; Kringle; 1.  
SMART; SM00181; EGF; 1.  
SMART; SM00130; KR; 1.  
SMART; SM00020; Tryp SPC; 1.  
PROSITE; PS00022; EGF\_1; 1.  
PROSITE; PS01186; EGF\_2; 1.  
PROSITE; PS00026; EGF\_3; 1.  
PROSITE; PS00021; KRINGLE\_1; 1.  
PROSITE; PS00070; KRINGLE\_2; 1.  
PROSITE; PS02040; TRYPSIN\_DOM; 1.  
PROSITE; PS00134; TRYPSIN\_HIS; 1.  
PROSITE; PS00135; TRYPSIN\_SER; 1.

KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
KW Kringle; EGF-like domain; Signal; Multigene family.  
FT SIGNAL 1 36  
FT CHAIN 37 431  
FT DOMAIN 37 75  
FT DOMAIN 82 163  
FT DOMAIN 179 431  
FT ACT\_SITE 226 226  
FT ACT\_SITE 275 275  
FT ACT\_SITE 382 382  
FT DISULFID 41 52  
FT DISULFID 46 63  
FT DISULFID 65 74  
FT DISULFID 82 163  
FT DISULFID 103 145  
FT DISULFID 134 158  
FT DISULFID 168 229  
FT DISULFID 211 227  
FT DISULFID 219 288  
FT DISULFID 313 388  
FT DISULFID 345 361  
FT DISULFID 378 406  
FT CARBOHYD 139 139  
FT CARBOHYD 352 352  
SQ SEQUENCE 431 AA; 48221 MW; 69958675B162CBF CRC64;  
Query Match 41.4%; Score 328.5; DB 1; Length 431;  
Best Local Similarity 47.6%; Pred. No. 5.3e-25;  
Matches 59; Conservative 16; Mismatches 48; Indels 1; Gaps 1  
QY 9 SNCDCLNGTCVSNKYFNSNIHWNCNPKKFGGHCIDKSKTCYEGNHFYRGKASTDTMG 68  
DB 42 SELRCFNGGTCWQAASFSDP-VCQCPKGTGTGQCEVDTHATCYKQGVYGTWSTSESG 100  
QY 69 RECLPWN SATVLQOTYHARRSDALQLGLGKHNHCNPNRRPWCYVQVGLKPLVQECV 128  
DB 101 AQCIWNNSLLTFTYNGRRSDAITLGLGHNHCNPNDRNSKPCWYVTKASKFILEFCV 160  
QY 129 HQCA 132  
DB 161 PVCS 164  
RESULT 10  
TPA\_HUMAN  
ID TPA\_HUMAN STANDARD; PRT; 562 AA.  
AC P00750; Q15103;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)  
DE (t-PA) (t-plasminogen activator) (Alteplase) (Retepase).  
GN PLAT.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Melanoma;  
RX MEDLINE=83115262; PubMed=6337343;  
RA Pennica D., Holmes W.E., Kohn W.J., Harkins R.N., Vohar G.A., Ward C.A., Bennett W.F., Yelverton E., Seeburg P.H., Heyneker H.L., Goeddel D.V., Collen D.;  
RT "Cloning and expression of human tissue-type plasminogen activator cDNA in *E. coli*.";  
RL Nature 301:214-221(1983).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=fetal lung;  
RX MEDLINE=88262579; PubMed=3133640;  
RA Sasaki H., Saito Y., Hayashi M., Otsuka K., Niwa M.;  
RT "Nucleotide sequence of the tissue-type plasminogen activator cDNA



RT from human fetal lung cells.";  
RL Nucleic Acids Res. 16:5695-5695(1988).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88054470; PubMed=2824147;  
RA Reddy V.B., Garramone A.J., Sasak H., Wei C.-M., Watkins P., Galli J.,  
RA Hsiung N.;  
RT "Expression of human uterine tissue-type plasminogen activator in  
RT mouse cells using BPV vectors.";  
RL DNA 6:461-472(1987).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86196143; PubMed=3009482;  
RA Priesner Degen S.J., Rajput B., Reich E.;  
RA "The human tissue plasminogen activator gene";  
RL J. Biol. Chem. 261:6972-6985(1986).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=84298137; PubMed=6089198;  
RA Ny T., Elgh F., Lund B.;  
RT "The structure of the human tissue-type plasminogen activator gene:  
RT correlation of intron and exon structures to functional and  
RT structural domains.";  
RL Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359(1984).  
RN [6]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86284200; PubMed=3090401;  
RA Harris T.J., Patel T., Marston F.A., Little S., Entage J.S.;  
RA Opdenakker G., Volckert G., Rombaux W., Billiau A., Somer P.;  
RT "Cloning of cDNA coding for human tissue-type plasminogen activator  
RT and its expression in Escherichia coli.";  
RL Mol. Biol. Med. 3:279-292(1986).  
RN [7]  
RP SEQUENCE FROM N.A. (ISOFORM SHORT).  
RC TISSUE=Umbilical vein;  
RX MEDLINE=90192129; PubMed=2107528;  
RA Siebert F.B., Fong K.;  
RT "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from  
RT human endothelial cells.";  
RL Nucleic Acids Res. 18:1086-1086(1990).  
RN [8]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=2238957; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udell T.B., Ioshizaki S., Carninci P., Frange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.D.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Mugny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [9]  
RP SEQUENCE OF 212-361 FROM N.A.  
RX MEDLINE=83169656; PubMed=6572897;  
RA Edlund T., Ny T., Raanby M., Heden L.-O., Palm G., Holmgren E.,  
RA Josephson S.;  
RT "Isolation of cDNA sequences coding for a part of human tissue  
RT plasminogen activator.";  
RL Proc. Natl. Acad. Sci. U.S.A. 80:349-352(1983).  
RN [10]

RP SEQUENCE OF 1-36 FROM N.A.  
RX MEDLINE=85289338; PubMed=3161893;  
RA Fisher R., Waller E.K., Grossi G., Thompson D., Tizard R.,  
RA Schleuning W.-D.;  
RT "Isolation and characterization of the human tissue-type plasminogen  
RT activator structural gene including its 5' flanking region.";  
RL J. Biol. Chem. 260:11223-11230(1985).  
RN [11]  
RP SEQUENCE OF 31-562 FROM N.A.  
RX MEDLINE=91291340; PubMed=1368681;  
RA Itagaki Y., Yasuda H., Morinaga T., Mitsuda S., Higashio K.;  
RT "Purification and characterization of tissue plasminogen activator  
RT secreted by human embryonic lung diploid fibroblasts, IMR-90 cells.";  
RL Agric. Biol. Chem. 55:1225-1232(1991).  
RN [12]  
RP SEQUENCE OF 36-562.  
RC TISSUE=Melanoma;  
RX MEDLINE=85000468; PubMed=6433976;  
RA Pohl G., Kaelinroem M., Bergsdorf N., Wallen P., Joernvall H.;  
RT "Tissue plasminogen activator: peptide analyses confirm an indirectly  
RT derived amino acid sequence, identify the active site serine residue,  
RT establish glycosylation sites, and localize variant differences.";  
RL Biochemistry 23:3701-3707(1984).  
RN [13]  
RP SEQUENCE OF 33-52 AND 311-330.  
RC TISSUE=Melanoma;  
RX MEDLINE=83209620; PubMed=6682760;  
RA Wallen P., Pohl G., Bergsdorf N., Raanby M., Ny T., Joernvall H.;  
RT "Purification and characterization of a melanoma cell plasminogen  
RT activator.";  
RL Eur. J. Biochem. 132:681-686(1983).  
RN [14]  
RP STRUCTURE OF CARBOHYDRATES.  
RX MEDLINE=90092112; PubMed=2513186;  
RA Pfeiffer G., Schmidt M., Strube K.-H., Geyer R.;  
RT "Carbohydrate structure of recombinant human uterine tissue  
RT plasminogen activator expressed in mouse epithelial cells.";  
RL Eur. J. Biochem. 186:273-286(1989).  
RN [15]  
RP CARBOHYDRATE-LINKAGE SITE THR-96.  
RX MEDLINE=91159408; PubMed=1900431;  
RA Harris R.J., Leonard C.K., Guzzetta A.W., Spellman M.W.;  
RT "Tissue plasminogen activator has an O-linked fucose attached to  
RT threonine-61 in the epidermal growth factor domain.";  
RL Biochemistry 30:2311-2314(1991).  
RN [16]  
RP DISULFIDE BONDS IN KRINGLE 2.  
RX MEDLINE=91244765; PubMed=1645336;  
RA Vlahos C.J., Wilhelm O.G., Kassell T., Jaekunas S.R., Bang N.U.;  
RT "Disulfide pairing of the recombinant kringle-2 domain of tissue  
RT plasminogen activator produced in Escherichia coli.";  
RL J. Biol. Chem. 266:10070-10072(1991).  
RN [17]  
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF CATALYTIC DOMAIN.  
RX MEDLINE=96200985; PubMed=8613982;  
RA Lamba D., Bauer M., Huber R., Fischer S., Rudolph R., Kohnert U.,  
RA Bode W.;  
RT "The 2.3 A crystal structure of the catalytic domain of recombinant  
RT two-chain human tissue-type plasminogen activator.";  
RL J. Mol. Biol. 258:117-135(1996).  
RN [18]  
RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF CATALYTIC DOMAIN.  
RX MEDLINE=97449126; PubMed=9305622;  
RA Renatus M., Engh R.A., Stubbs M.T., Huber R., Fischer S., Kohnert U.,  
RA Bode W.;  
RT "Lysine 156 promotes the anomalous proenzyme activity of tPA: X-ray  
RT crystal structure of single-chain human tPA.";  
RL EMBO J. 16:4797-4805(1997).  
RN [19]  
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF KRINGLE 2.  
RX MEDLINE=92119803; PubMed=1310033;  
RA de Vos A., Ultsch M.H., Kelley R.F., Padmanabhan K., Tulinsky A.,  
RA Westbrook M.L., Kossiakof A.A.;

RT "Crystal structure of the kringle 2 domain of tissue plasminogen  
RT activator at 2.4-A resolution.";  
RL Biochemistry 31:270-279(1992).  
RN [20]  
RP STRUCTURE BY NMR OF KRINGLE 2.  
RA MEDLINE=90122799; PubMed=2558718;  
RA Byeon I.-J.L., Kelley R.F., Llinas M.;  
RT "1H NMR structural characterization of a recombinant kringle 2 domain  
RT from human tissue-type plasminogen activator.";  
RL Biochemistry 28:9350-9360(1989).  
RN [21]  
RP STRUCTURE BY NMR OF KRINGLE 2.  
RX MEDLINE=92106329; PubMed=1762144;  
RA Byeon I.-J.L., Llinas M.;  
RA "Solution structure of the tissue-type plasminogen activator kringle  
RT 2 domain complexed to 6-aminohexanoic acid an antifibrinolytic  
RT drug.";  
RL J. Mol. Biol. 222:1035-1051(1991).  
RN [23]  
Query Match 41.4%; Score 328.5; DB 1; Length 562;  
Best Local Similarity 46.3%; Pred. No. 6.9e-25;  
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;  
QY 3 ELHQVP-SNCD---CLNGTGVSNKYPSNIHWNCNPKKFGGHCIDKSKTCYSGNGHFY 58  
DB 77 QCHSVFVKSCEPRFCNGTCQOALYFSDF-VQCPEGPAKCEIDTRATCYSDQGISY 135  
QY 59 RKGASTDTMGRLCPWNSATVLQQTTHAHRSDALQGLGKKNYCRNPNRRRRCYVQVG 118  
DB 136 RGTWSTAESGABCTWNSSALAKPKYSGRRPDALRLGLGNHNYCRNPNRDRSKPKCYVFK 195  
QY 119 LKPLVQECVHDCAG 134  
DB 196 GKYSEFCSTPACSEG 211  
RESULT 11  
TPA RAT STANDARD; PRT; 559 AA.  
AC P19837;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Tissue-type plasminogen activator precursor (EC 3.4.21.66) (tPA)  
DE (t-PA) (t-plasminogen activator).  
GN PLAT.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
[1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89170114; PubMed=3148445;  
RA Ny T.; Leonardson G.; Hsueh A.J.W.;  
RT "Cloning and characterization of a cDNA for rat tissue-type  
RT plasminogen activator.";  
RL DNA 7:671-677(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90130448; PubMed=2105315;  
RA Feng P.; Ohlsson M.; Ny T.;  
RT "The structure of the TARA-less rat tissue-type plasminogen activator  
RT gene. Species-specific sequence divergences in the promoter predict  
RT differences in regulation of gene expression.";  
RL J. Biol. Chem. 265:2022-2027(1990).

CC -1- FUNCTION: Converts the abundant, but inactive, zymogen plasminogen  
CC to plasmin by hydrolyzing a single Arg-Val bond in plasminogen. By  
CC controlling plasmin-mediated proteolysis, it plays an important  
CC role in tissue remodeling and degradation, in cell migration and  
CC many other physiopathological events.  
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
CC plasminogen to form plasmin.  
CC -1- SUBUNIT: Heterodimer of chain A and chain B held by a disulfide  
CC bond.  
CC -1- SUBCELLULAR LOCATION: Secreted; extracellular.  
CC -1- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER  
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER  
CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.  
CC -1- MISCELLANEOUS: Binds to the kringle structure of the fibrin A  
CC chain. Binding to fibrin enhances its catalytic activity.  
CC -1- SIMILARITY: Belongs to peptidase family S1.  
CC -1- SIMILARITY: Contains 1 EGF-like domain.  
CC -1- SIMILARITY: Contains 1 fibronectin type I domain.  
CC -1- SIMILARITY: Contains 2 kringle domains.  
CC -----  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboratio  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on it  
CC use by non-profit institutions as long as its content is in no wa  
CC modified and this statement is not removed. Usage by and for commercia  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M23697; AAA41812.1; -;  
CC EMBL; M31197; AAA42261.1; -;  
CC EMBL; M31185; AAA42261.1; JOINED.  
CC EMBL; M31186; AAA42261.1; JOINED.  
CC EMBL; M31187; AAA42261.1; JOINED.  
CC EMBL; M31188; AAA42261.1; JOINED.  
CC EMBL; M31189; AAA42261.1; JOINED.  
CC EMBL; M31190; AAA42261.1; JOINED.  
CC EMBL; M31191; AAA42261.1; JOINED.  
CC EMBL; M31192; AAA42261.1; JOINED.  
CC EMBL; M31193; AAA42261.1; JOINED.  
CC EMBL; M31194; AAA42261.1; JOINED.  
CC EMBL; M31195; AAA42261.1; JOINED.  
CC EMBL; M31196; AAA42261.1; JOINED.  
CC EMBL; A19619; CAA01482.1; -;  
CC PIR; A35029; A35029.  
CC HSP; P00750; IRTF.  
CC MEROPS; S01.232; -;  
CC InterPro; IPR009003; Cys\_Ser\_trypsin.  
CC InterPro; IPR006209; EGF\_like.  
CC InterPro; IPR000083; Fibrinctnl.  
CC InterPro; IPR006210; IEGF.  
CC InterPro; IPR000001; Kringle.  
CC InterPro; IPR001254; Peptidase\_S1.  
CC InterPro; IPR001314; Peptidase\_S1A.  
CC Pfam; PF00008; EGF; 1.  
CC Pfam; PF00039; fnl; 1.  
CC Pfam; PF00051; Kringle; 2.  
CC Pfam; PF00089; trypsin; 1.  
CC PRINTS; PR00722; CHYMOTRYPSIN.  
CC PRINTS; PR00018; KRINGLE.  
CC ProDom; PD000395; Kringle; 2.  
CC SMART; SM00181; EGF; 1.  
CC SMART; SM00058; FNI; 1.  
CC SMART; SM00130; KR; 2.  
CC SMART; SMC0020; Tryp\_Spc; 1.  
CC PROSITE; PS00022; EGF\_1; 1.  
CC PROSITE; PS01186; EGF\_2; 1.  
CC PROSITE; PS50026; EGF\_3; 1.  
CC PROSITE; PS01253; FIBRONECTIN\_1; 1.  
CC PROSITE; PS00021; KRINGLE\_1; 2.  
CC PROSITE; PS50070; KRINGLE\_2; 2.  
CC PROSITE; PS50240; TRYPSIN\_DOM; 1.  
CC PROSITE; PS00134; TRYPSIN\_HIS; 1.  
CC PROSITE; PS00135; TRYPSIN\_SER; 1.

KW Plasminogen activator; Hydrolyase; Serine protease; Glycoprotein;  
KW SIGNAL; Kringle; EGF-like domain; Repeat; Signal.  
FT PROPEP 18 29  
FT CHAIN 30 559  
FT CHAIN 30 308  
FT CHAIN 309 559  
FT DOMAIN 36 78  
FT DOMAIN 79 117  
FT DOMAIN 124 205  
FT DOMAIN 213 294  
FT DOMAIN 309 559  
FT ACT\_SITE 355 355  
FT ACT\_SITE 404 404  
FT ACT\_SITE 510 510  
FT DISULFID 38 68  
FT DISULFID 66 75  
FT DISULFID 83 94  
FT DISULFID 107 116  
FT DISULFID 124 205  
FT DISULFID 145 187  
FT DISULFID 176 200  
FT DISULFID 213 294  
FT DISULFID 234 276  
FT DISULFID 265 299  
FT DISULFID 297 428  
FT DISULFID 340 356  
FT DISULFID 348 417  
FT DISULFID 442 516  
FT DISULFID 474 490  
FT DISULFID 506 534  
FT CARBOHYD 149 149  
FT CARBOHYD 481 481  
FT CONFLICT 380 380  
SQ SEQUENCE 559 AA; 62903 MD; 7DBD3809C1D1C921 CRC64;  
Query Match 40.5%; Score 321.5; DB 1; Length 559;  
Best Local Similarity 45.6%; Pred. No. 3.3e-24;  
Matches 62; Conservative 14; Mismatches 55; Indels 5; Gaps 2;  
QY 3 ELHQVP-----SNCDCLNGGTGVSNKYFVNIHWCNCPKFGGQHCIDKSKTCYENGHFY 58  
DB 74 QCHSVPRVRCSEPRFCNGTTCQALYFSDP-VQCPDGVGRCRDIDFRAICFEGQGIT 132  
QY 59 RGKASTDTMGRCPLPNSATVLTQTYHAHRSDALQLGLKHNCRPNRRRPWCYVQVG 118  
DB 133 RGTWSTAENGAECINWSSALSQKPYARRNAIKLGLGNHNYCRNPDVDKPCYVFK 192  
QY 119 LKPLVQECMHDCADG 134  
DB 193 GKVTTFCTPACFKG 208  
RESULT 12  
TPA\_MOUSE  
ID TPA\_MOUSE STANDARD; PRT; 559 AA.  
AC P11214; Q91VP2;  
DT 01-JUL-1989 (Rel. 11, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)  
DE (t-PA) (t-plasminogen activator).  
GN PLAT  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88087303; PubMed=2826484;

RA Rickles R.J., Darrow A.L., Strickland S.;  
RT "Molecular cloning of complementary DNA to mouse tissue plasminogen  
RT activator mRNA and its expression during F9 teratocarcinoma cell  
RT differentiation.";  
RL J. Biol. Chem. 263:1563-1569(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Mammary Gland;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.N., Schuler G.D.,  
RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hest L.,  
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Falley J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smallos D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -|- FUNCTION: Converts the abundant, but inactive, zymogen plasminogen  
CC to plasmin by hydrolyzing a single Arg-Val bond in plasminogen. By  
CC controlling plasmin-mediated proteolysis, it plays an important  
CC role in tissue remodeling and degradation, in cell migration and  
CC many other physiological events.  
CC -|- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
CC plasminogen to form plasmin.  
CC -|- SUBUNIT: Heterodimer of chain A and chain B held by a disulfide  
CC bond.  
CC -|- SUBCELLULAR LOCATION: Secreted; extracellular.  
CC -|- FM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER  
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER  
CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.  
CC -|- MISCELLANEOUS: Binds to the kringle structure of the fibrin A  
CC chain. Binding to fibrin enhances its catalytic activity.  
CC -|- SIMILARITY: Belongs to peptidase family S1.  
CC -|- SIMILARITY: Contains 1 EGF-like domain.  
CC -|- SIMILARITY: Contains 1 fibronectin type I domain.  
CC -|- SIMILARITY: Contains 2 kringle domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collabora-  
CC tion between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on it  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercia-  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; J03520; AAA40470.1; -;  
CC EMBL; BC011256; AAH11256.1; -;  
CC PIR; A29941; A29941.  
CC HSSP; P00750; 1A5H.  
CC MEROPS; S01.232; -;  
CC MGD; MGI:97610; Plat.  
CC InterPro; IPR003003; Cys Ser trypsin.  
CC InterPro; IPR006209; EGF like.  
CC InterPro; IPR000083; Fibrinctn1.  
CC InterPro; IPR006210; LEGF.  
CC InterPro; IPR000001; Kringle.  
CC InterPro; IPR001254; Peptidase\_S1.  
CC InterPro; IPR001314; Peptidase\_S1A.  
CC Pfam; PF00008; EGF; 1.  
CC Pfam; PF00039; fn1; 1.  
CC Pfam; PF00051; Kringle; 2.  
CC -----

DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PRO0722; CHYMOTRYPSIN.  
 DR PRINTS; PRO0018; KRINGLE.  
 DR PRODOM; PD000395; Kringle; 2.  
 DR SMART; SM00181; EGF; 1.  
 DR SMART; SM00058; FN1; 1.  
 DR SMART; SM00130; KR; 2.  
 DR SMART; SM00020; TYP SPc; 1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS00026; EGF\_3; 1.  
 DR PROSITE; PS01253; FIBRONECTIN\_1; 1.  
 DR PROSITE; PS00021; KRINGLE\_1; 2.  
 DR PROSITE; PS00070; KRINGLE\_2; 2.  
 DR PROSITE; PS00240; TRYPsin DOM; 1.  
 DR PROSITE; PS00134; TRYPsin HIS; 1.  
 DR PROSITE; PS00135; TRYPsin SER; 1.  
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
 KW Plasma; Kringle; EGF-like domain; Repeat; Signal.  
 FT SIGNAL 1 17  
 FT PROPEP 18 29  
 FT CHAIN 30 559  
 FT CHAIN 30 308  
 FT CHAIN 309 559  
 FT DOMAIN 36 78  
 FT DOMAIN 79 117  
 FT DOMAIN 124 205  
 FT DOMAIN 213 294  
 FT DOMAIN 309 559  
 FT ACT\_SITE 355 355  
 FT ACT\_SITE 404 404  
 FT ACT\_SITE 510 510  
 FT DISULFID 38 68  
 FT DISULFID 66 75  
 FT DISULFID 83 94  
 FT DISULFID 88 105  
 FT DISULFID 107 116  
 FT DISULFID 124 205  
 FT DISULFID 145 187  
 FT DISULFID 176 200  
 FT DISULFID 213 294  
 FT DISULFID 234 276  
 FT DISULFID 265 289  
 FT DISULFID 297 428  
 FT DISULFID 340 356  
 FT DISULFID 348 417  
 FT DISULFID 442 516  
 FT DISULFID 474 490  
 FT DISULFID 506 534  
 FT CARBOHYD 149 149  
 FT CARBOHYD 481 481  
 FT CONFLICT 260 260  
 FT CONFLICT 325 325  
 SQ SEQUENCE 559 AA; 63122 MW; 8CCEE2BDB94514D9 CRC64;  
 Query Match 39.8%; Score 315.5; DB 1; Length 559;  
 Best Local Similarity 44.5%; Pred. No. 1.3e-23;  
 Matches 61; Conservative 15; Mismatches 56; Indels 5; Gaps 2;  
 QY 3-ELHQP-----SNCDLNGGTGVSNGKFSNIHNCNPKKPGGQHCIDSKICYENGHIFY 58  
 DB 74 QCHSVFVRSCSPRCFCNGGTCQALYSDF-VQCQPDGVGKRCIDIDTRATCFEEQGIT 132  
 QY 59 RGKASDTWTGRPCLPWNSATVLQOTYHAHRSALQGLGKHNCRNPNRRRPWCYQVG 118  
 DB 133 RGTWSTAEGACINNSVLSKPNRPNRAIKLGLGNHNYCRNPNEDLKPWCYVFK 192  
 QY 119 LKPLVOECMVHDCADGK 135  
 DB 193 GKYTEFCSTPACPCKGK 209

# RESULT 13

URL DESRO STANDARD; PRT; 477 AA.  
 AC P98119;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Salivary plasminogen activator alpha 1 precursor (BC 3.4.21.68) (DSPA  
 DE alpha-1).  
 OS Desmodus rotundus (Vampire bat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;  
 OC Desmodontinae; Desmodus.  
 OX NCBI\_TaxID=9430;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Salivary Gland;  
 RX MEDLINE=92039036; PubMed=1937019;  
 RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,  
 RA Alagon A., Donner P., Schleuning W.D.;  
 RT "The plasminogen activator family from the salivary gland of the  
 RT vampire bat Desmodus rotundus: cloning and expression.";  
 RL Gene 105:229-237(1991).  
 RN [2]  
 RP CHARACTERIZATION.  
 RX MEDLINE=93393059; PubMed=1309059;  
 RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,  
 RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,  
 RA Donner P.;  
 RT "Plasminogen activators from the saliva of Desmodus rotundus (common  
 RT vampire bat): unique fibrin specificity.";  
 RL Ann. N.Y. Acad. Sci. 667:395-403(1992).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
 RC TISSUE=Salivary Gland;  
 RX MEDLINE=98022741; PubMed=9354616;  
 RA Renatus M., Stubbs M.T., Huber R., Bringmann P., Donner P.,  
 RA Schleuning W.D., Bode W.;  
 RT "Catalytic domain structure of vampire bat plasminogen activator: a  
 RT molecular paradigm for proteolysis without activation cleavage.";  
 RL Biochemistry 36:13483-13493(1997).  
 CC -!- FUNCTION: Probably essential to support the feeding habits of this  
 CC exclusively haematophagous animal. Potent thrombolytic agent.  
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
 CC plasminogen to form plasmin.  
 CC -!- ENZYME REGULATION: Activity toward plasminogen is stimulated in  
 CC the presence of fibrin I.  
 CC -!- SUBUNIT: Monomer.  
 CC -!- DOMAIN: The fibronectin type-I domain mediates binding to fibrin,  
 CC and the kringle domain apparently mediates fibrin-induced  
 CC stimulation of activity.  
 CC -!- SIMILARITY: Belongs to peptidase family S1.  
 CC -!- SIMILARITY: Contains 1 EGF-like domain.  
 CC -!- SIMILARITY: Contains 1 fibronectin type I domain.  
 CC -!- SIMILARITY: Contains 1 kringle domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboratio  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on it  
 CC use by non-profit institutions as long as its content is in no wa  
 CC modified and this statement is not removed. Usage by and for commercia  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M63987; AAA31591.1; -;  
 DR EMBL; M63986; AAA31592.1; -;  
 DR PIR; JS0597; JS0597.  
 DR PDB; 1A51; 23-MAR-99.  
 DR MEROPS; S01.232; -;  
 DR GlycoSuiteDB; P98119; -;  
 DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
 DR InterPro; IPR006209; EGF\_like.

DR InterPro: IPR000083; Fibrinctn1.  
DR InterPro: IPR006210; IEGF.  
DR InterPro: IPR000001; Kringle.  
DR InterPro: IPR001254; Peptidase\_S1.  
DR InterPro: IPR001314; Peptidase\_S1A.  
DR Pfam: PF00008; EGF\_1.  
DR Pfam: PF00039; fn1; 1.  
DR Pfam: PF00051; kringle; 1.  
DR Pfam: PF00089; tyrosin; 1.  
DR PRINTS: PR00722; CHYMOTRYPSIN.  
DR PRINTS: PD00018; KRINGLE.  
DR ProDom: PD000395; Kringle; 1.  
DR SMART: SM00181; EGF; 1.  
DR SMART: SM00058; FN1; 1.  
DR SMART: SM00130; KR; 1.  
DR SMART: SM00020; TYP\_SPC; 1.  
DR PROSITE: PS00022; EGF\_1; 1.  
DR PROSITE: PS01186; EGF\_2; 1.  
DR PROSITE: PS00026; EGF\_3; 1.  
DR PROSITE: PS01253; FIBRONECTIN\_1; 1.  
DR PROSITE: PS00021; KRINGLE\_1; 1.  
DR PROSITE: PS00070; KRINGLE\_2; 1.  
DR PROSITE: PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE: PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
KW Plasminogen activation; Hydrolase, Serine protease; Glycoprotein;  
KW Kringle; EGF-like domain; Signal; Multigene family; 3D-structure.  
FT SIGNAL 1 36  
FT CHAIN 37 477  
FT DOMAIN 40 82  
FT DOMAIN 83 121  
FT DOMAIN 128 209  
FT DOMAIN 225 477  
FT ACT SITE 272 272  
FT ACT SITE 321 321  
FT ACT SITE 428 428  
FT DISULFID 42 72  
FT DISULFID 70 79  
FT DISULFID 87 98  
FT DISULFID 92 109  
FT DISULFID 111 120  
FT DISULFID 128 209  
FT DISULFID 149 191  
FT DISULFID 180 204  
FT DISULFID 214 345  
FT DISULFID 257 273  
FT DISULFID 265 334  
FT DISULFID 359 434  
FT DISULFID 391 407  
FT DISULFID 424 452  
FT CARBOHYD 153 153  
FT CARBOHYD 398 398  
FT TURN 214 215  
FT STRAND 223 224  
FT TURN 226 227  
FT STRAND 230 231  
FT HELIX 234 236  
FT TURN 238 239  
FT STRAND 240 243  
FT STRAND 254 263  
FT TURN 264 265  
FT STRAND 266 269  
FT HELIX 271 273  
FT TURN 280 282  
FT STRAND 284 287  
FT TURN 297 298  
FT STRAND 300 309  
FT TURN 311 312  
FT TURN 315 317  
FT TURN 319 320  
FT STRAND 323 328

FT STRAND 338 338  
FT TURN 339 340  
FT STRAND 341 341  
FT STRAND 345 345  
FT TURN 349 350  
FT TURN 355 356  
FT STRAND 358 363  
FT STRAND 366 366  
FT STRAND 374 374  
FT STRAND 379 385  
FT STRAND 388 390  
FT HELIX 393 398  
FT TURN 403 404  
FT STRAND 405 409  
FT TURN 425 426  
FT TURN 428 429  
FT STRAND 431 436  
FT TURN 437 438  
FT STRAND 439 448  
FT TURN 455 456  
FT STRAND 459 463  
FT STRAND 464 467  
FT HELIX 468 474  
SQ SEQUENCE 477 AA; 53616 MW; AA06FDI739C10E5E CRC64;  
Query Match 39.2%; Score 310.5; DB 1; Length 477;  
Best Local Similarity 45.5%; Pred. NO. 3.4e-23;  
Matches 60; Conservative 14; Mismatches 53; Indels 5; Gaps 3  
QY 5 HQVPSN-CD---CLNGGTCVSNKYFSNIHWNCNPKKFGQCHCEIDKSKTCYEGNGHFYRG 60  
DB 80 HTVPVNSCSEPRCFNGGTCWQAVYFSDF-VCCPAGYTGKRCVDTRATCYEGQGVYRG 138  
QY 61 KASTDTMGRLPWSNATVLOQTYHAHRSDALQLGKKNYCNPNRRPHCYVOGLK 120  
DB 139 TWSTAESRVEICINWNSLLTRTYNGRMPDAPNLGLGNHNYCRNPNRGPAPWCYVIKAGK 198  
QY 121 PLVQECWMDCA 132  
DB 199 FTSESCSVFVCS 210  
RESULT 14  
TPA\_BOVIN  
ID TPA\_BOVIN STANDARD; PRT; 566 AA.  
AC Q28198; (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)  
DE (t-PA) (t-plasminogen activator).  
GN PLAT.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RA Ravn P., Berglund L., Petersen T.E.;  
RT "Cloning and characterization of the bovine plasminogen activators uPA  
and tPA";  
RL int. Dairy J. 5:605-617(1995).  
CC -!- FUNCTION: Converts the abundant, but inactive, zymogen plasminogen  
to plasmin by hydrolyzing a single Arg-Val bond in plasminogen. By  
controlling plasmin-mediated proteolysis, it plays an important  
role in tissue remodeling and degradation, in cell migration and  
many other physiological events.  
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
plasminogen to form plasmin.  
CC -!- SUBUNIT: Heterodimer of chain A and chain B held by a disulfide  
bond.

CC -!- SUBCELLULAR LOCATION: Secreted; extracellular.  
CC -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER  
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER  
CC ARG-314 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.  
CC -!- MISCELLANEOUS: Binds to the kringle structure of the fibrin A  
CC chain. Binding to fibrin enhances its catalytic activity.  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
CC -!- SIMILARITY: Contains 1 EGF-like domain.  
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.  
CC -!- SIMILARITY: Contains 2 kringle domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X85800; CAA59795.1; --  
CC HSSP: P00750; IRTF.  
CC MEROPS: S01.232; --  
CC InterPro: IPR009003; Cys\_Ser\_trypsin.  
CC InterPro: IPR006209; EGF-like.  
CC InterPro: IPR000083; Fibrinctn1.  
CC InterPro: IPR006210; IEGF.  
CC InterPro: IPR000001; Kringle.  
CC InterPro: IPR001254; Peptidase\_S1.  
CC InterPro: IPR001314; Peptidase\_S1A.  
CC Pfam: PF00008; EGF; 1.  
CC Pfam: PF00039; fn1; 1.  
CC Pfam: PF00051; kringle; 2.  
CC Pfam: PF00089; trypsin; 1.  
CC PRINTS: PR00722; CHYMOTRYPSIN.  
CC PRINTS: PR00018; KRINGLE.  
CC ProDom: PD000395; Kringle; 2.  
CC SMART: SM00181; EGF; 1.  
CC SMART: SM00058; FN1; 1.  
CC SMART: SM00130; KR; 2.  
CC SMART: SM00020; Tryp\_SPC; 1.  
CC PROSITE: PS00022; EGF\_1; 1.  
CC PROSITE: PS01186; EGF\_2; 1.  
CC PROSITE: PS00026; EGF\_3; 1.  
CC PROSITE: PS01253; FIBRONECTIN\_1; 1.  
CC PROSITE: PS00021; KRINGLE\_1; 1.  
CC PROSITE: PS00070; KRINGLE\_2; 2.  
CC PROSITE: PS00240; TRYPsin\_DOM; 1.  
CC PROSITE: PS00134; TRYPsin\_HIS; 1.  
CC PROSITE: PS00135; TRYPsin\_SER; 1.  
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
KW Plasma; Kringle; EGF-like domain; Repeat; Signal.  
FT SIGNAL 1 21  
FT PROPEP 22 33  
FT CHAIN 34 566  
FT CHAIN 34 314  
FT CHAIN 315 566  
FT CHAIN 40 82  
FT DOMAIN 83 121  
FT DOMAIN 128 209  
FT DOMAIN 219 300  
FT DOMAIN 315 566  
FT ACT\_SITE 361 361  
FT ACT\_SITE 410 410  
FT ACT\_SITE 517 517  
FT DISULFID 42 72  
FT DISULFID 70 79  
FT DISULFID 87 98  
FT DISULFID 92 109  
FT DISULFID 111 120  
FT DISULFID 128 209  
FT DISULFID 149 191

FT DISULFID 180 204 BY SIMILARITY.  
FT DISULFID 219 300 BY SIMILARITY.  
FT DISULFID 282 295 BY SIMILARITY.  
FT DISULFID 271 295 BY SIMILARITY.  
FT DISULFID 303 434 INTERCHAIN (BY SIMILARITY).  
FT DISULFID 346 362 BY SIMILARITY.  
FT DISULFID 354 423 BY SIMILARITY.  
FT DISULFID 448 523 BY SIMILARITY.  
FT DISULFID 480 496 BY SIMILARITY.  
FT DISULFID 513 541 BY SIMILARITY.  
FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 487 487 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 566 AA; 63701 MW; 2E86BBE4E32276C3 CRC64;  
Query Match 37.9%; Score 300.5; DB 1; Length 566;  
Best Local Similarity 44.0%; Pred. No. 3.9e-22;  
Matches 59; Conservative 16; Mismatches 54; Indels 5; Gaps 3  
QY 3 ELHQVP-SNCD---CLNGTCVSNKYFSNIHWCNCPKFGGQHCIDSKSTCYEGNGHY 58  
DB 78 QCHSVPVRSCEPWCFCNGTCTQALYSDF-VCCQPEGFMGKLCIDATATCYKQGVAY 136  
QY 59 RGVASTDTMGRCLPWNSTATVQQTYHAHRSALQGLGKHYCRNPDRPWCYVQVG 118  
DB 137 RGTWSTAESGAECANWNSGLAMKTPYSGRRPNARLGLGNHNYCRNPQDQSKPWCYFKA 196  
QY 119 LKPLVOECMVHDC A 132  
DB 197 GKIYSEFCSTPACA 210  
RESULT 15  
HGFA HUMAN STANDARD; PRT; 655 AA.  
AC Q04756; Q14726;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF  
DN activator) (HGFA).  
GN HGFA.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Liver, and Serum;  
RX MEDLINE=93252878; PubMed=7683665;  
RA Miyakawa K.; Shimomura T.; Kitamura A.; Kondo J.; Morimoto Y.,  
RA Kitamura N.;  
RT "Molecular cloning and sequence analysis of the cDNA for a human  
RT serine protease responsible for activation of hepatocyte growth  
RT factor. Structural similarity of the protease precursor to blood  
RT coagulation factor XII.";  
RL J. Biol. Chem. 268:10024-10028 (1993).  
RN [2]  
RP SEQUENCE OF 40-655 FROM N.A.  
RA Zhao S., Odell C.;  
RA Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Activates hepatocyte growth factor (HGF) by  
CC converting it from a single chain to a heterodimeric form.  
CC -!- SUBUNIT: Dimer of a short chain and a long chain linked by a  
CC disulfide bond.  
CC -!- SUBCELLULAR LOCATION: Secreted as an inactive single-chain  
CC precursor and is then activated to a heterodimeric form.  
CC -!- TISSUE SPECIFICITY: Liver.  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
CC -!- SIMILARITY: Contains 2 EGF-like domains.  
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.  
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.  
CC -!- SIMILARITY: Contains 1 kringle domain.  
CC -!- CAUTION: It is uncertain whether Met-1 is the initiator.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: D14012; BAA03113.1; -  
 DR EMBL: Z69923; -; NOT\_ANNOTATED\_CDS.  
 DR FIR: A46688; A46688.  
 DR HSP: P00763; LDPO.  
 DR MEROPS: S01.228; -  
 DR Genew: HGNC:4894; HGFAC.  
 DR MIM: 604552; -  
 DR GO: GO:0005576; C:extracellular; TAS.  
 DR GO: GO:0004252; F:serine-type endopeptidase activity; TAS.  
 DR GO: GO:0008508; P:proteolysis and peptidolysis; TAS.  
 DR InterPro: IPR009003; Cys Ser trypsin.  
 DR InterPro: IPR000742; EGF\_2.  
 DR InterPro: IPR006209; EGF-like.  
 DR InterPro: IPR000083; Fibnctn1.  
 DR InterPro: IPR000562; FN\_Type\_II.  
 DR InterPro: IPR006210; IEGF.  
 DR InterPro: IPR000001; Kringle.  
 DR InterPro: IPR001254; Peptidase\_S1.  
 DR InterPro: IPR001314; Peptidase\_S1A.  
 DR Pfam: PF00008; EGF\_2.  
 DR Pfam: PF00039; fn1; 1.  
 DR Pfam: PF00040; fn2; 1.  
 DR Pfam: PF00051; kringle; 1.  
 DR Pfam: PF00089; trypsin; 1.  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR PRINTS: PR00013; FNTYPEII.  
 DR PRINTS: PR00018; KRINGLE.  
 DR ProDom: PD000995; FN\_Type\_II; 1.  
 DR ProDom: PD000395; Kringle; 1.  
 DR SMART: SM00181; EGF; 2.  
 DR SMART: SM00358; FN1; 1.  
 DR SMART: SM00559; FN2; 1.  
 DR SMART: SM00130; KR; 1.  
 DR SMART: SM00020; TRYP\_SPC; 1.  
 DR PROSITE: PS00022; EGF\_1; 2.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 DR PROSITE: PS50026; EGF\_3; 2.  
 DR PROSITE: PS01253; FIBRONECTIN\_1; 1.  
 DR PROSITE: PS00023; FIBRONECTIN\_2; 1.  
 DR PROSITE: PS00021; KRINGLE\_1; 1.  
 DR PROSITE: PS00070; KRINGLE\_2; 1.  
 DR PROSITE: PS50240; TRYPsin\_DOM; 1.  
 DR PROSITE: PS00134; TRYPsin\_HIS; 1.  
 DR PROSITE: PS00135; TRYPsin\_SER; 1.  
 KW Hydrolase; glycoprotein; Plasma; Serine protease; Kringle; Signal;  
 KW EGF-like domain; Repeat; Zymogen.  
 FT SIGNAL 1 30  
 FT PROPEP 31 372  
 FT CHAIN 373 407  
 FT CHAIN 408 655  
 FT CHAIN 108 148  
 FT DOMAIN 160 198  
 FT DOMAIN 200 240  
 FT DOMAIN 241 279  
 FT DOMAIN 286 367  
 FT DOMAIN 408 655  
 FT ACT\_SITE 447 447  
 FT ACT\_SITE 497 497  
 FT ACT\_SITE 598 598  
 FT ACT\_SITE 108 133  
 FT DISULFID 122 148  
 FT DISULFID 164 175

FT DISULFID 169 186 BY SIMILARITY.  
 FT DISULFID 188 197 BY SIMILARITY.  
 FT DISULFID 202 230 BY SIMILARITY.  
 FT DISULFID 228 237 BY SIMILARITY.  
 FT DISULFID 245 256 BY SIMILARITY.  
 FT DISULFID 250 267 BY SIMILARITY.  
 FT DISULFID 269 278 BY SIMILARITY.  
 FT DISULFID 286 367 BY SIMILARITY.  
 FT DISULFID 307 349 BY SIMILARITY.  
 FT DISULFID 338 362 BY SIMILARITY.  
 FT DISULFID 394 521 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 432 448 BY SIMILARITY.  
 FT DISULFID 440 510 BY SIMILARITY.  
 FT DISULFID 535 604 BY SIMILARITY.  
 FT DISULFID 567 583 BY SIMILARITY.  
 FT DISULFID 594 622 BY SIMILARITY.  
 FT CARBOHYD 48 48 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 468 468 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 492 492 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 546 546 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 644 644 R -> Q (IN REF. 2).  
 SQ SEQUENCE 655 AA; 70681 MW; 2CF72F1E1B862ED7 CRC64;  
 Query Match 32.8%; Score 260.5; DB 1; Length 655;  
 Best Local Similarity 43.7%; Pred. No. 3.8e-18;  
 Matches 52; Conservative 8; Mismatches 42; Indels 17; Gaps 2  
 QY 5 HOVPSNCDCLNGGTCVSNKYFSNIHW-----CNCPKKFGQGHCEIDKSKTCYRGNGH 56  
 Db 242 HTACLSPCLNGGTC-----HLIVATGTTVCACPPGCFAGRLCNIEPDERCFLNGT 292  
 QY 57 FYRGKASTDTMGRCPLPWSATVLQTYHAHRSDALQLGLGKHNYCNPENRRPWCYV 115  
 Db 293 GYRGVASTSASGLSCLAWNSDLLYQELHVDVSGAALLGLGPHAYCRNPENRRPWCYV 351  
 RESULT 16  
 FA12\_CAVPO STANDARD; PRT; 603 AA.  
 AC Q04962;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)  
 DE (HAF) (Fragment).  
 GN F12  
 OS Cavia porcellus (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.  
 OC NCBI\_TaxID=10141;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-37; 318-332 AND 359-373.  
 RC TISSUE=Liver;  
 RX MEDLINE=93003367; PubMed=1390917;  
 RA Semba U., Yamamoto T., Kunisada T., Shibuya Y., Tanase S.,  
 RA Kanbara T., Okabe H.;  
 RT "Primary structure of guinea-pig Hageman factor: sequence around the  
 RT cleavage site differs from the human molecule.";  
 RL Biochim. Biophys. Acta 1159:113-121(1992).  
 CC -!- FUNCTION: Factor XII is a serum glycoprotein that participates in  
 CC the initiation of blood coagulation, fibrinolysis, and the  
 CC generation of bradykinin and angiotensin.  
 CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-Ile bonds in factor  
 CC VII to form factor VIIa and factor XI to form factor Xia.  
 CC -!- MISCELLANEOUS: Factor XII, prekallikrein, and HMW kininogen form a  
 CC complex bound to an anionic surface. Prekallikrein is cleaved by  
 CC factor XII to form kallikrein, which then cleaves factor XII first  
 CC to alpha-factor XIIa and then to beta-factor Xia. Alpha-factor  
 CC Xia activates factor XI to factor Xia.  
 CC -!- SIMILARITY: Belongs to peptidase family S1.  
 CC -!- SIMILARITY: Contains 2 EGF-like domains.  
 CC -!- SIMILARITY: Contains 1 fibronectin type I domain.





the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```
EMBL; AF099017; AAF02489.1; -
EMBL; AF224724; AAF34712.1; -
HSSP; P00763; IDPO.
MEROPS; S01.228; -.
MGD; MGI:1859281; Hgfac.
InterPro; IPR009003; Cys_Ser_trypsin.
InterPro; IPR000742; EGF_2.
InterPro; IPR006209; EGF-like.
InterPro; IPR000083; Fibronctni.
InterPro; IPR000862; FN_functni.
InterPro; IPR000862; FN_Type_II.
InterPro; IPR006210; IEGF.
InterPro; IPR000001; Kringle.
InterPro; IPR001254; Peptidase_S1.
InterPro; IPR001314; Peptidase_S1A.
Pfam; PF00008; EGF; 2.
Pfam; PF00039; fcl; 1.
Pfam; PF00040; fcl2; 1.
Pfam; PF00051; kringle; 1.
PRINTS; PRO0089; trypsin; 1.
PRINTS; PRO0722; CHYMOTRYPSIN.
PRINTS; PRO0013; ENTPEII.
PRINTS; PRO0018; KRINGLE.
ProDom; PD000995; FN_Type_II; 1.
SMART; SM00181; EGF; 2.
SMART; SM000395; Kringle; 1.
SMART; SM00059; FN2; 1.
SMART; SM00130; KR; 1.
SMART; SM00020; Tryp_Sp; 1.
PROSITE; PS00022; EGF_1; 2.
PROSITE; PS01186; EGF_2; 1.
PROSITE; PS00026; EGF_3; 2.
PROSITE; PS01253; FIBRONECTIN_1; 1.
PROSITE; PS00023; FIBRONECTIN_2; 1.
PROSITE; PS00021; KRINGLE_1; 1.
PROSITE; PS00070; KRINGLE_2; 1.
PROSITE; PS00240; TRYPSIN_DOM; 1.
PROSITE; PS00134; TRYPSIN_HIS; 1.
PROSITE; PS00135; TRYPSIN_SER; 1.
Hydrolase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
EGF-like domain; Repeat; Zymogen.
FT SIGNAL 1 29 BY SIMILARITY.
FT PROPEP 30 369 CLEAVED IN ACTIVE FORM (BY SIMILARITY).
FT CHAIN 370 405 HEPATOCYTE GROWTH FACTOR ACTIVATOR SHORT
CHAIN.
FT CHAIN 406 653 HEPATOCYTE GROWTH FACTOR ACTIVATOR LONG
CHAIN.
FT DOMAIN 105 145 FIBRONECTIN TYPE-II.
FT DOMAIN 157 195 EGF-LIKE 1.
FT DOMAIN 197 237 FIBRONECTIN TYPE-I.
FT DOMAIN 238 276 EGF-LIKE 2.
FT DOMAIN 283 364 KRINGLE.
FT DOMAIN 406 653 SERINE PROTEASE.
FT ACT_SITE 445 445 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 495 495 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 596 596 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 105 130 BY SIMILARITY.
FT DISULFID 119 145 BY SIMILARITY.
FT DISULFID 161 172 BY SIMILARITY.
FT DISULFID 165 183 BY SIMILARITY.
FT DISULFID 185 194 BY SIMILARITY.
FT DISULFID 199 227 BY SIMILARITY.
FT DISULFID 225 234 BY SIMILARITY.
FT DISULFID 242 253 BY SIMILARITY.
FT DISULFID 247 264 BY SIMILARITY.
FT DISULFID 266 275 BY SIMILARITY.
FT DISULFID 283 364 BY SIMILARITY.
FT DISULFID 304 346 BY SIMILARITY.
```

```
FT DISULFID 335 359 BY SIMILARITY.
FT DISULFID 392 519 INTERCHAIN (BY SIMILARITY).
FT DISULFID 430 446 BY SIMILARITY.
FT DISULFID 438 508 BY SIMILARITY.
FT DISULFID 533 602 BY SIMILARITY.
FT DISULFID 565 581 BY SIMILARITY.
FT DISULFID 592 620 BY SIMILARITY.
FT CARBOHYD 39 39 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 47 47 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 63 63 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 466 466 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 544 544 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 164 164 G -> W (IN REF. 2).
SQ SEQUENCE 653 AA; 70567 MW; 88B4B2055DF7DC CRC64;

Query Match 31.6%; Score 250.5; DB 1; Length 653;
Best Local Similarity 42.0%; Pred. No. 3.6e-17;
Matches 50; Conservative 11; Mismatches 41; Indels 17; Gaps 2

QY 5 HQVPNSCDLNGTCTVSNKYFSNIHW-----CNCPEKFGGQCEIDKSKTCEYGNHG 56
Db 239 HTACLSPPCLNGTCTC-----HLVGTGTSVCTCLPGYAGRECNIVPTIEHCLNGT 289

QY 57 FYRGKASTDTMRPCLPWSATVLOQTVHAHSDALQLGLGHNYCRPNDRRRPWCYV 115
Db 290 EYRGVASTAASGLSCLAWNSDLLYQELHVDVSAVAALLGLGPHAYCRNPDKDEREWCYV 348

RESULT 18
FA12_HUMAN STANDARD; PRT; 615 AA.
ID P00748; P78339;
AC P00748; P78339;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
DE (HAF).
GN F12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88007593; PubMed=2888762;
RA Cool D.E., McGillivray R.T.A.;
RT "Characterization of the human blood coagulation factor XII gene.
RT Intron/exon gene organization and analysis of the 5'-flanking
RT region."
RL J. Biol. Chem. 262:13662-13673(1987).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANTS ALA-207; ASP-545 AND HIS-605.
RA Rieder M.J., Armel T.Z., Carrington D.P., Ozuna M., Kuldane S.A.,
RA Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 4-615 FROM N.A.
RX MEDLINE=86176794; PubMed=3754331;
RA Tripodi M., Citarella F., Guida S., Galeffi P., Fantoni A.,
RA Cortese R.;
RT "cDNA sequence coding for human coagulation factor XII (Hageman).";
RL Nucleic Acids Res. 14:3146-3146(1986).
RN [4]
RP SEQUENCE OF 14-615 FROM N.A.
RX MEDLINE=8603830; PubMed=3877053;
RA Cool D.E., Edgell C.-J.S., Louie G.V., Zoller M.J., Brayer G.D.,
RA McGillivray R.T.A.;
RT "Characterization of human blood coagulation factor XII cDNA.
RT Prediction of the primary structure of factor XII and the tertiary
RT structure of beta-factor XIIa."
RL J. Biol. Chem. 260:13666-13676(1985).
RN [5]
```

RP SEQUENCE OF 146-615 FROM N.A.  
RX MEDLINE=86216049; PubMed=3011063;  
RA Que B.G., Davie B.W.;  
RT "Characterization of a cDNA coding for human factor XII (Hageman factor).";  
RL Biochemistry 25:1525-1528(1986).  
RN [6]  
RP SEQUENCE OF 20-379  
RX MEDLINE=85182674; PubMed=3886654;  
RA McMullen B.A., Fujikawa K.;  
RT "Amino acid sequence of the heavy chain of human alpha-factor XIIa (activated Hageman factor).";  
RL J. Biol. Chem. 260:5328-5341(1985).  
RN [7]  
RP SEQUENCE OF 354-362 AND 373-615  
RX MEDLINE=83291041; PubMed=6604055;  
RA Fujikawa K., McMullen B.A.;  
RT "Amino acid sequence of human beta-factor XIIa.";  
RL J. Biol. Chem. 258:10924-10933(1983).  
RN [8]  
RP SEQUENCE OF 561-615 FROM N.A.  
RC TISSUE=Blood;  
RX MEDLINE=96133302; PubMed=8528215;  
RA Schloesser M., Hofferbert S., Bartz U., Lutze G., Lammle B., Engel W.;  
RT "The novel acceptor splice site mutation 11396(G-->A) in the factor XII gene causes a truncated transcript in cross-reacting material RT negative patients.";  
RL Hum. Mol. Genet. 4:1235-1237(1995).  
RN [9]  
RP CARBOHYDRATE-LINKAGE SITE THR-109  
RX MEDLINE=92184750; PubMed=1544894;  
RA Harris R.J., Ling V.T., Spellman M.W.;  
RT "O-linked fucose is present in the first epidermal growth factor domain of factor XII but not protein C.";  
RL J. Biol. Chem. 267:5102-5107(1992).  
RN [10]  
RP VARIANT WASHINGTON D.C. SER-590  
RX MEDLINE=90046788; PubMed=2510163;  
RA Miyata T., Kawabata S.-I., Iwanaga S., Takahashi I., Alving B., Saito H.;  
RT "Coagulation factor XII (Hageman factor) Washington D.C.: inactive factor XIIa results from Cys-571-->Ser substitution.";  
RL Proc. Natl. Acad. Sci. U.S.A. 86:8319-8322(1989).  
RN [11]  
RP VARIANT LOCARNO PRO-372  
RX MEDLINE=94325559; PubMed=8049433;  
RA Hovinga J.K., Schaller J., Stricker H., Wuillemin W.A., Furian M., Laemmle B.;  
RT "Coagulation factor XII Locarno: the functional defect is caused by the amino acid substitution Arg-353-->Pro leading to loss of a kallikrein cleavage site.";  
RL Blood 84:1173-1181(1994).  
RN [12]  
RP VARIANT TENRI CYS-53  
RX MEDLINE=99290785; PubMed=103161128;  
RA Kondo S., Tokunaga F., Kawano S., Oono Y., Kumagai S., Koide T.;  
RT "Factor XII Tenri, a novel cross-reacting material negative factor XII deficiency, occurs through a protease-mediated degradation.";  
RL Blood 93:4300-4308(1999).  
CC -!- FUNCTION: Factor XII is a serum glycoprotein that participates in the initiation of blood coagulation, fibrinolysis, and the generation of bradykinin and angiotensin.  
CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Ile bonds in factor VII to form factor VIIa and factor XI to form factor Xla.  
CC -!- PTM: O- AND N-GLYCOSYLATED.  
CC -!- DISEASE: Defects in F12 do not cause any clinical symptoms. The sole effect is that whole-blood clotting time is prolonged.  
CC -!- MISCELLANEOUS: Factor XII, prekallikrein, and HMW kininogen form a complex bound to an anionic surface. Prekallikrein is cleaved by factor XII to form kallikrein, which then cleaves factor XII first to alpha-factor XIIa and then to beta-factor XIIa. Alpha-factor XIIa activates factor XI to factor Xla.  
CC -!- SIMILARITY: Belongs to peptidase family S1.

CC -!- SIMILARITY: Contains 2 EGF-like domains.  
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.  
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.  
CC -!- SIMILARITY: Contains 1 kringle domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC ENBL; M1315; AAA70235.1; --  
CC ENBL; AFS38691; AAM97932.1; --  
CC ENBL; M1723; AAA51986.1; --  
CC ENBL; M17466; AAB59490.1; --  
CC ENBL; M17464; AAB59490.1; JOINED.  
CC ENBL; M17465; AAB59490.1; JOINED.  
CC ENBL; M13147; AAA70224.1; --  
CC ENBL; U71274; AAB51203.1; --  
CC PIR; A29411; KFHJ12.  
CC HSSP; P00763; IDPO.  
CC MEROPS; S01.211; --  
CC Genew; HGNC:3530; F12.  
CC MIM; 234000; --  
CC GO; GO:0003805; F.blood coagulation factor XI activity; TAS.  
CC GO; GO:0003806; F.blood coagulation factor XII activity; TAS.  
CC GO; GO:0008236; F.blood coagulation; TAS.  
CC GO; GO:0007596; F.blood coagulation; TAS.  
CC InterPro; IPR009003; Cys\_Ser\_trypsin.  
CC InterPro; IPR006209; EGF\_like.  
CC InterPro; IPR000083; Fibrinctn1.  
CC InterPro; IPR000562; FN\_Type\_II.  
CC InterPro; IPR006210; IEGF.  
CC InterPro; IPR000001; Kringle.  
CC InterPro; IPR001254; Peptidase\_S1.  
CC InterPro; IPR001314; Peptidase\_S1A.  
CC Pfam; PF00008; EGF; 2.  
CC Pfam; PF00039; fn1; 1.  
CC Pfam; PF00040; fn2; 1.  
CC Pfam; PF00051; kringle; 1.  
CC Pfam; PF00089; trypsin; 1.  
CC PRINTS; PR00722; CHYMOTRYPSIN.  
CC PRINTS; PR00013; ENTPEPII.  
CC PRINTS; PR00018; KRINGLE.  
CC ProDom; PD000995; FN\_Type\_II; 1.  
CC ProDom; PD000395; Kringle; 1.  
CC SMART; SM00181; EGF; 2.  
CC SMART; SM00058; FN1; 1.  
CC SMART; SM00059; FN2; 1.  
CC SMART; SM00130; KR; 1.  
CC SMART; SM00020; Tryp\_Spc; 1.  
CC PROSITE; PS00022; EGF\_1; 2.  
CC PROSITE; PS01186; EGF\_2; 1.  
CC PROSITE; PS50036; EGF\_3; 2.  
CC PROSITE; PS01253; FIBRONECTIN\_1; 1.  
CC PROSITE; PS00023; FIBRONECTIN\_2; 1.  
CC PROSITE; PS00021; KRINGLE\_1; 1.  
CC PROSITE; PS50070; KRINGLE\_2; 1.  
CC PROSITE; PS50240; TRYPSIN\_DOM; 1.  
CC PROSITE; PS00134; TRYPSIN\_HIS; 1.  
CC PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Glycoprotein; blood coagulation; Plasma; Kringle; Serine protease;  
KW Hydrolyase; Fibrinolysis; Signal; EGF-like domain; Repeat; Zymogen;  
KW Polymorphism; Disease mutation.  
FT SIGNAL 1 19  
FT CHAIN 20 372 ALPHA-FACTOR XIIA HEAVY CHAIN.  
FT CHAIN 373 615 BETA-FACTOR XIIA PART 1.  
FT CHAIN 354 362 BETA-FACTOR XIIA PART 2.  
FT CHAIN 373 615 FIBRONECTIN TYPE-II.  
FT DOMAIN 47 88 EGF-LIKE 1.  
FT DOMAIN 94 131

FT DOMAIN 133 173 FIBRONECTIN TYPE-1.  
FT DOMAIN 174 210 EGF-LIKE 2.  
FT DOMAIN 217 295 KRINGLE.  
FT DOMAIN 236 349 PRO-RICH.  
FT DOMAIN 373 615 SERINE PROTEASE.  
FT CARBOHYD 109 109 O-LINKED (FUC).  
FT CARBOHYD 249 249 N-LINKED (GLCNAC...).  
FT CARBOHYD 299 299 O-LINKED (POTENTIAL).  
FT CARBOHYD 305 305 O-LINKED (POTENTIAL).  
FT CARBOHYD 308 308 O-LINKED (POTENTIAL).  
FT CARBOHYD 328 328 O-LINKED (POTENTIAL).  
FT CARBOHYD 329 329 O-LINKED (POTENTIAL).  
  
Query Match 29.4%; Score 233; DB 1; Length 615;  
Best Local Similarity 42.3%; Pred. No. 1.8e-15;  
Matches 44; Conservative 12; Mismatches 40; Indels 8; Gaps 3;  
  
QY 13 CLINGTCVSNKFSNIHWCNCPKFGGHCEDSKCTVEGNGHYRGKASTDTMGRECL 72  
DB 183 CLHGRCLE---VEGRLCHCPVGTGPFCDVDTKASYDGRGLSYRLATTTLSGAPCQ 239  
  
QY 73 PMSATVLQOQTY-HAHRSDALQGLGKHNCRPNRRPWCYV 115  
DB 240 PWAS-----EATYRNVTAEQARNWGLGHAFCRNPDIRPWCYV 279  
  
RESULT 19  
URTG\_DESGO STANDARD; PRT; 394 AA.  
AC P49150;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Salivary plasminogen activator gamma precursor (EC 3.4.21.68) (DSPA Gamma).  
DE Desmodus rotundus (Vampire bat).  
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;  
OC Desmodontinae; Desmodus.  
OX NCBI\_TaxID=9430;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Salivary gland;  
RX MEDLINE=92039036; PubMed=1937019;  
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P., Alagon A., Donner P., Schleuning W.D.;  
RT "The plasminogen activator family from the salivary gland of the vampire bat Desmodus rotundus: cloning and expression."  
RL Gene 105:229-237(1991).  
RN [2]  
RP CHARACTERIZATION.  
RX MEDLINE=9339059; PubMed=1309059;  
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T., Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W., Donner P.;  
RT "Plasminogen activators from the saliva of Desmodus rotundus (common vampire bat): unique fibrin specificity."  
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).  
CC -!- FUNCTION: Probably essential to support the feeding habits of this exclusively haematophagous animal. Probable potent thrombolytic agent.  
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.  
CC -!- SUBUNIT: Monomer.  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
CC -!- SIMILARITY: Contains 1 kringle domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC -----  
CC EMBL; M63990; AAA31595.1; --  
CC PIR; JS0600; JS0600.  
CC HSSP; P98119; 1A51.  
CC MEROPS; S01.239; --  
CC InterPro; IPR009003; Cys\_Ser\_trypsin.  
CC InterPro; IPR000001; Kringle.  
CC InterPro; IPR001254; Peptidase\_S1.  
CC InterPro; IPR001314; Peptidase\_S1A.  
CC Pfam; PF00051; kringle; 1.  
CC Pfam; PF00089; trypsin; 1.  
CC PRINTS; PR00722; CHYMOTRYPSIN.  
CC PRINTS; PR00018; KRINGLE.  
CC ProDom; PD000395; Kringle; 1.  
CC SMART; SM00130; KR; 1.  
CC SMART; SM00020; Tryp\_SPC; 1.  
CC PROSITE; PS00021; KRINGLE 1; 1.  
CC PROSITE; PS00070; KRINGLE 2; 1.  
CC PROSITE; PS00240; TRYPSIN\_DOM; 1.  
CC PROSITE; PS00134; TRYPSIN\_HIS; 1.  
CC PROSITE; PS00135; TRYPSIN\_SER; 1.  
CC Plasminogen activation; Hydrolase; Serine protease; Glycoprotein; Kringle; Signal; Multigene family.  
KW SIGNAL 1 36 POTENTIAL.  
FT CHAIN 37 394 SALIVARY PLASMINOGEN ACTIVATOR GAMMA.  
FT DOMAIN 45 126 KRINGLE.  
FT DOMAIN 142 394 SERINE PROTEASE.  
FT ACT\_SITE 189 189 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 238 238 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 345 345 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT DISULFID 45 126 BY SIMILARITY.  
FT DISULFID 66 108 BY SIMILARITY.  
FT DISULFID 97 121 BY SIMILARITY.  
FT DISULFID 131 262 BY SIMILARITY.  
FT DISULFID 174 190 BY SIMILARITY.  
FT DISULFID 182 251 BY SIMILARITY.  
FT DISULFID 276 351 BY SIMILARITY.  
FT DISULFID 308 324 BY SIMILARITY.  
FT DISULFID 341 369 BY SIMILARITY.  
FT CARBOHYD 315 315 N-LINKED (GLCNAC... (POTENTIAL).  
SQ SEQUENCE 394 AA; 44105 MW; 9CCD6F52F3D81FCD CRC64;  
  
Query Match 28.1%; Score 223; DB 1; Length 394;  
Best Local Similarity 44.3%; Pred. No. 1.1e-14;  
Matches 39; Conservative 13; Mismatches 36; Indels 0; Gaps 0  
  
QY 45 DKSKTVEGNGHFYRGKASTDTMGRCPLPMSATVLQOQTYHAHRSDALQGLGKHNCRN 104  
DB 40 DPHATCYKQGVYRGWTSISGACQINWNSLLIERTYNGRMPKGLGKHNCRN 99  
  
QY 105 PDNRRPWCYVQVGLKPLVQECMVHDC 132  
DB 100 PDGASKPCYVIRAKFTSESCSVVPCS 127  
  
RESULT 20  
FA12\_BOVIN STANDARD; PRT; 593 AA.  
AC P98140;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)  
DE (HAP) (fragment).  
GN F12.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.

RC TISSUE=Liver;  
RX MEDLINE=94242782; PubMed=8186251;  
RA Shibuya Y., Samba U., Okabe H., Kambara T., Yamamoto T.;  
RT "Primary structure of bovine Hageman factor (blood coagulation factor XII): comparison with human and guinea pig molecules."; Biochim. Biophys. Acta 1206:63-70(1994).  
RN [2].  
RP SEQUENCE OF 10-21; 350-364 AND 525-550.  
RX MEDLINE=7182112; PubMed=861210;  
RA Fujikawa K., Walsh A.K., Davis W.B.;  
RT "Isolation and characterization of bovine factor XII (Hageman factor).";  
RN Biochemistry 16:2270-2278(1977).  
CC !- FUNCTION: Factor XII is a serum glycoprotein that participates in the initiation of blood coagulation, fibrinolysis, and the generation of bradykinin and angiotensin.  
CC !- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Ile bonds in factor VII to form factor VIIa and factor XI to form factor XIIa.  
CC !- PTM: O- AND N-GLYCOSYLATED (BY SIMILARITY).  
CC !- MISCELLANEOUS: Factor XII, prekallikrein, and HMW kininogen form a complex bound to an anionic surface. Prekallikrein is cleaved by factor XII to form kallikrein, which then cleaves factor XI first to alpha-factor XIIa and then to beta-factor XIIa. Alpha-factor XIIa activates factor XI to factor Xia. Bovine factor XII is cleaved only to alpha-factor XIIa as it lacks the trypsin/kallikrein cleavage site.  
CC !- SIMILARITY: Belongs to peptidase family S1.  
CC !- SIMILARITY: Contains 2 EGF-like domains.  
CC !- SIMILARITY: Contains 1 fibronectin type I domain.  
CC !- SIMILARITY: Contains 1 fibronectin type II domain.  
CC !- SIMILARITY: Contains 1 kringle domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
DR EMBL; S70164; AAB30804.2; --  
DR PIR; S45281; S45281.  
DR HSSP; P00763; 1DPO.  
DR MEROPS; S01.211; --  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR000083; Fibrinctn1.  
DR InterPro; IPR000562; FN Type\_II.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00008; EGF; 2.  
DR Pfam; PF00040; fn2; 1.  
DR Pfam; PF00051; kringle; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00013; FNTYPEII.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000995; FN Type II; 1.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00181; EGF; 2.  
DR SMART; SM00058; FN1; 1.  
DR SMART; SM00059; FN2; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS00022; EGF\_1; 2.  
DR PROSITE; PS01186; EGF\_2; 2.  
DR PROSITE; PS00026; EGF\_3; 2.  
DR PROSITE; PS01253; FIBONECTIN\_1; 1.  
DR PROSITE; PS00023; FIBONECTIN\_2; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.

DR PROSITE; PS00070; KRINGLE\_2; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;  
KW Hydrolase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.  
FT NON\_TER 1  
FT SIGNAL <1 9  
FT CHAIN 10 349  
FT CHAIN 350 593  
FT DOMAIN 37 78  
FT DOMAIN 84 121  
FT DOMAIN 123 163  
FT DOMAIN 164 200  
FT DOMAIN 207 287  
FT DOMAIN 297 333  
FT DOMAIN 350 593  
FT ACT\_SITE 389 399  
FT ACT\_SITE 438 438  
FT ACT\_SITE 541 541  
FT DISULFID 88 100  
FT DISULFID 94 109  
FT DISULFID 111 120  
FT DISULFID 125 153  
FT DISULFID 151 160  
FT DISULFID 168 179  
FT DISULFID 173 188  
FT DISULFID 190 199  
FT DISULFID 207 287  
FT DISULFID 230 269  
FT DISULFID 258 282  
FT DISULFID 336 463  
FT DISULFID 374 390  
FT DISULFID 382 452  
FT DISULFID 413 416  
FT DISULFID 479 547  
FT DISULFID 510 526  
FT DISULFID 537 568  
FT CARBOHYD 99 99  
FT CARBOHYD 241 241  
FT CARBOHYD 263 263  
FT CARBOHYD 410 410  
SQ SEQUENCE 593 AA; 65148 MW; 721592BA792BD61F CRC64;  
Query Match 27.4%; Score 217; DB 1; Length 593;  
Best Local Similarity 35.7%; Pred. No. 6.3e-14;  
Matches 46; Conservative 15; Mismatches 58; Indels 10; Gaps 4  
Qy 6 QVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGCHCEIDKSKTCYE--GNGHFYRGKAS 63  
Db 166 QVCRTPCLNGSDSLQAE---GHLRCRCAPSFAGRLCDVLDLKASCYDDRRDGLSYRGWAG 222  
Qy 64 TDTWGRPCLPWNSATVLQOQTY-HAHRSDALQLGLGHKNTYCRNPDNRPRPCYVQVGLKPL 122  
Db 223 TLTSGAPCQSWAS-----EATYWNVTAEQVLNWLGLGDHAFGRNPDNDRPNCFTWKGRDLS 278  
Qy 123 VQECWVHDC 131  
Db 279 WNYCLAPC 287  
RESULT 21  
RCR1 DROME  
ID \_RCR1 DROME STANDARD; PRT; 685 AA.  
AC Q24438;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Tyrosine-protein kinase transmembrane receptor Ror precursor  
DE (EC 2.7.1.112) (drom).  
GN ROR OR CG4926.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

[illegible]

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement ([see http://www.isb-sib.ch/announce](http://www.isb-sib.ch/announce)) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

-----  
CC  
DR EMBL; L20297; AAA28860.1; --  
DR EMBL; A5003628; AAP28865.1; --  
DR EMBL; AA002908; CAAU5743.1; --  
DR PIR; A48289; A48289.  
DR HSP; F11362; IFCK.  
DR FlyBase; FBgn010407; Ror.  
DR GO; GO:0016021; C:integral to membrane; NAS.  
DR GO; GO:0004713; P:protein-tyrosine kinase activity; NAS.  
DR GO; GO:0007417; P:central nervous system development; IEP.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; NAS.  
DR InterPro; IPRO00024; Fz domain.  
DR InterPro; IPRO00001; Kringle.  
DR InterPro; IPRO00719; Prot\_kinase.  
DR InterPro; IPRO02011; ReceptTyKinsII.  
DR InterPro; IPRO01245; Tyr\_pkinase.  
DR InterPro; IPRO08266; Tyr\_pkinase\_AS.  
DR Pfam; PF00051; kringle\_1;  
DR Pfam; PF00069; pkinase\_1;  
DR PRINTS; PRO0018; KRINGLE.  
DR PRINTS; PRO0109; TYRKINASE.  
DR Prodom; PD000395; Kringle; 1.  
DR Prodom; PD000001; Prot\_Kinase; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00219; TyRkC; 1.  
DR PROSITE; PS00038; FZ; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS00070; KRINGLE\_2; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS00239; RECEPTOR\_TYR\_KIN\_I; 1.  
DR Transferrase; Kinase; Tyrosine-protein kinase; ATP-binding; Receptor;  
KW Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation;  
KW Developmental protein.  
FT SIGNAL 1..24  
FT CHAIN 25..685  
  
FT FT DOMAIN 25 317  
FT TRANSMEM 318 338  
FT DOMAIN 339 685  
FT DOMAIN 36 225  
FT DOMAIN 236 310  
FT DOMAIN 410 677  
FT NP\_BIND 416 424  
FT BINDING 442 442  
FT ACT\_SITE 539 539  
FT MOD\_RES 565 565  
  
FT FT MOD\_RES 569 569  
FT FT MOD\_RES 570 570  
FT CARBOHYD 45 45  
FT CARBOHYD 63 63  
FT CARBOHYD 129 129  
FT CARBOHYD 144 144  
FT CARBOHYD 250 250  
SEQUENCE 685 AA; 78142 MW; N26162D27D5FD7C7 CRC64;  
  
Query Match 21.3%; Score 169; DB 1; Length 685;  
Best Local Similarity 28.8%; Pred. No. 3.7e-09;  
Matches 38; Conservative 27; Mismatches 35; Indels 32; Gaps 62

Db 212 DQKLPQHQCISLGIIT-----EVDTNCTWEGSTYRGVA 244

Oy 63 STDTMGRCLPWN SATVLQQTYHAHRSDALQLGLGHKNYCRPN-D--RRSPWCYVQVGLKP 121  
:  
Db 250 NVSAGKPCFLRW---SWLMKEI-----SDFPEL-IGQ-NYCRNPGSVENSPWCFVDSSRER 300  
:  
Oy 122 LVQECCWHDCAD 133  
::: | : |||  
Db 301 IIELCDIFPKAD 312  
:  
RESULT 22  
APOA\_HUMAN  
ID APOA\_HUMAN STANDARD; PRT; 4548 AA.  
AC P08519.  
DT 01-AUG-1988 (Rel. 08, Created)  
DI 01-AUG-1988 (Rel. 08, last sequence update)  
DE 10-OCT-2003 (Rel. 42, last annotation update)  
DE Apolipoprotein(a) precursor (EC 3.4.21.-) (Apo(a)) (Lp(a)).  
GN LPA.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88039109; PubMed=3670400;  
RA McLean J.W., Tomlinson J.E., Kuang W.-J., Eaton D.L., Chen E.Y.,  
RA Fless G.M., Scanu A.M., Lawn R.M.;  
RA "cDNA sequence of human apolipoprotein(a) is homologous to  
RT plasminogen." ;  
RL Nature 330:132-137(1987).  
RN [2]  
RP SERINE PROTEASE ACTIVITY.  
RX MEDLINE=90076123; PubMed=2531657;  
RA Salonen E.-M., Jaahainen M., Zardi L., Vaheri A., Ehrenholm C.;  
RA "Lipoprotein(a) binds to fibronectin and has serine proteinase  
RT activity capable of cleaving it." ;  
RL EMBO J. 8:4035-4040(1989).  
RN [3]  
RP REVIEW.  
RX MEDLINE=90049223; PubMed=2530631;  
RA Utermann G.;  
RL "The mysteries of lipoprotein(a)." ;  
RL Science 246:904-910(1989).  
RN [4]  
RP CHARACTERIZATION OF THE N- AND O-LINKED GLYCANS.  
RX MEDLINE=21130359; PubMed=11294842;  
RA Garner B., Merry A.H., Royle L., Harvey D.J., Rudd P.M., Thillet J.;  
RA "Structural elucidation of the N- and O-glycans of human  
RT apolipoprotein(a): role of o-glycans in conferring protease  
RT resistance." ;  
RL J. Biol. Chem. 276:22200-22208(2001).  
RN [5]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 4121-4208.  
RX MEDLINE=96217891; PubMed=8642595;  
RA Mikol V., Lograsso P.V., Boettcher B.R.;  
RT "Crystal structures of apolipoprotein(a) kringle IV37 free and  
RT complexed with 6-aminoheptanoic acid and with p-aminomethylbenzoic  
RT acid: existence of novel and expected binding modes." ;  
RL J. Mol. Biol. 256:751-761(1996).  
RN [6]  
RP VARIANT ARG-4193.  
RX MEDLINE=95002201; PubMed=7918682;  
RA Scanu A.M., Praffinger D., Lee J.C., Hinman J.;  
RA "A single point mutation (Trp72-->Arg) in human apo(a) kringle 4-37  
RT associated with a lysine binding defect in Lp(a)." ;  
RL Biochim. Biophys. Acta 1227:41-45(1994).  
RN [7]  
RP FUNCTION: Apo(a) is the main constituent of lipoprotein(a)  
CC (lp(a)). It has serine proteinase activity and is able of  
CC autoproteolysis. Inhibits tissue-type plasminogen activator 1.  
CC Lp(a) may be a ligand for megalin/Gp 330.  
CC -! SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and

-|- PMN: N- and O-glycosylated. The N-glycans are complex biantennary structures present in either a mono- or disialylated state. The O-glycans are mostly (80%) represented by the monosialylated core type I structure, NeuNAcalpha2-3Galbeta1-3GalNAc, with smaller amounts of disialylated and non-sialylated O-glycans also detected.

-|- DISEASE: Elevated plasma concentrations of apo(a) and its naturally occurring proteolytic fragments are correlated with atherosclerosis. Homology with plasminogen kringle IV and V is thought to underlie the atherogenicity of the protein, because the fragments are competing with plasminogen for fibrin(ogen) binding.

-|- MISCELLANEOUS: Apo(a) is known to be proteolytically cleaved, leading to the formation of the so called mini-Lp(a). Apo(a) fragments accumulate in atherosclerotic lesions, where they may promote thrombogenesis. O-glycosylation may limit the extent of proteolytic fragmentation.

-|- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily. Contains 38 kringle domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation. The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; X06290; CAA29618.1; --  
PIR; S00657; S00657.  
PDB; 1I71; 13-JUN-01.  
PDB; 1JFN; 28-JUN-02.  
PDB; 1KIV; 18-MAY-99.  
PDB; 3KIV; 18-MAY-99.  
PDB; 4KIV; 18-MAY-99.  
MEROPS; S01.226; --  
GeneID; HGNC:6667; LPA.  
MIM; 152200; --  
GO; GO:0004866; F:endopeptidase inhibitor activity; TAS.  
GO; GO:0008015; P:circulation; TAS.  
GO; GO:0009405; P:pathogenesis; TAS.  
InterPro; IPRO09003; Cys\_Sex\_trypsin.  
InterPro; IPRO00001; Kringle.  
InterPro; IPRO01254; Peptidase\_S1.  
InterPro; IPRO01314; Peptidase\_S1A.  
Pfam; PF00051; kringle; 38.  
Pfam; PF00089; trypsin; 1.  
PRINTS; PR00722; CHYMOTRYPSIN.  
PRINTS; FR00018; KRINGLE.  
ProDom; PD000395; Kringle; 38.  
SMART; SM00130; KR; 38.  
SMART; SM00020; Tryp\_Spc; 1.  
PROSITE; PS00021; KRINGLE\_1; 38.  
PROSITE; PS00070; KRINGLE\_2; 38.  
PROSITE; PS02040; TRYPSIN\_DOM; 1.  
PROSITE; PS00134; TRYPSIN\_HIS; 1.  
PROSITE; PS00133; TRYPSIN\_SER; 1.  
Hydrolase; Serine protease; Lipid transport; Plasma; Glycoprotein; Kringle; Repeat; Atherosclerosis; Signal; Polymorphism; 3D-structure.

SIGNAL	1	19
CHAIN	20	4548
FT	20	APOLIPOPROTEIN(A).
DOMAIN	131	KRINGLE TYPE IV, 1.
FT	244	KRINGLE TYPE IV, 2.
DOMAIN	245	KRINGLE TYPE IV, 3.
FT	359	KRINGLE TYPE IV, 4.
DOMAIN	473	KRINGLE TYPE IV, 5.
FT	586	KRINGLE TYPE IV, 6.
DOMAIN	597	KRINGLE TYPE IV, 7.
FT	701	KRINGLE TYPE IV, 8.
DOMAIN	815	KRINGLE TYPE IV, 9.
FT	929	KRINGLE TYPE IV, 10.
DOMAIN	1043	KRINGLE TYPE IV, 11.
FT	1156	KRINGLE TYPE IV, 12.
DOMAIN	1271	KRINGLE TYPE IV, 13.

```
FT DOMAIN 1385 1498 KRINGLE TYPE IV, 13.
FT DOMAIN 1499 1612 KRINGLE TYPE IV, 14.
FT DOMAIN 1613 1726 KRINGLE TYPE IV, 15.
FT DOMAIN 1727 1840 KRINGLE TYPE IV, 16.
FT DOMAIN 1841 1954 KRINGLE TYPE IV, 17.
FT DOMAIN 1955 2068 KRINGLE TYPE IV, 18.
FT DOMAIN 2069 2182 KRINGLE TYPE IV, 19.
FT DOMAIN 2183 2296 KRINGLE TYPE IV, 20.
FT DOMAIN 2297 2410 KRINGLE TYPE IV, 21.
FT DOMAIN 2411 2524 KRINGLE TYPE IV, 22.
FT DOMAIN 2525 2638 KRINGLE TYPE IV, 23.
FT DOMAIN 2639 2752 KRINGLE TYPE IV, 24.
FT DOMAIN 2753 2866 KRINGLE TYPE IV, 25.
FT DOMAIN 2867 2980 KRINGLE TYPE IV, 26.
FT DOMAIN 2981 3094 KRINGLE TYPE IV, 27.
FT DOMAIN 3095 3208 KRINGLE TYPE IV, 28.
FT DOMAIN 3209 3322 KRINGLE TYPE IV, 29.
FT DOMAIN 3323 3436 KRINGLE TYPE IV, 30.
FT DOMAIN 3437 3550 KRINGLE TYPE IV, 31.
FT DOMAIN 3551 3664 KRINGLE TYPE IV, 32.
FT DOMAIN 3665 3778 KRINGLE TYPE IV, 33.
FT DOMAIN 3779 3892 KRINGLE TYPE IV, 34.
FT DOMAIN 3893 4006 KRINGLE TYPE IV, 35.
FT DOMAIN 4007 4120 KRINGLE TYPE IV, 36.
FT DOMAIN 4121 4234 KRINGLE TYPE IV, 37.
FT DOMAIN 4235 4348 SERINE PROTEASE.
FT ACT SITE 4349 4462 CHARGE RELAY SYSTEM.
FT ACT SITE 4463 4576 CHARGE RELAY SYSTEM.
FT ACT SITE 4577 4690 CHARGE RELAY SYSTEM.
FT VARIANT 4691 4804 W -> R (LOSS OF LYSINE-SEPHAROSE BINDING).
FT SEQUENCE 4805 4918 /FTIC=VAR 006633.
FT SEQUENCE 4919 5032 501313 MW; 96921B296A465C5F CRC64;
Query Match 20.5%; Score 162.5; DB 1; Length 4548;
Best Local Similarity 33.1%; Pred. No. 1e-07; 46; Indels 25; Gaps 7;
Matches 41; Conservative 12; Mismatches 12;
QY 27 NIHW--CN---CP-----KFGGQHCIEDKST---CYEGNGHYRGKASTDTMGR 69
DB 3742 NVREWCNLTQCFVTSESSVATSTAVSEQAPTSPTQDCVHGDSYRGSGFSFTVTGR 3801
QY 70 PCLPWSATVLOQTYHAHRSALQGLGHKHYCRNPNRRPWCYVQGLKPLV--QECM 127
DB 3802 TCQSSSMT---PHWHQRTTEYPNGGLTRNCRNPDARIRPWCYT---MDPSVRWEYCN 3855
QY 128 VHDC 131
DB 3856 LTQC 3859
RESULT 23
KRM1_RAT
AC Q92454; STANDARD; PRT; 473 AA.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Kremen protein 1 precursor (kringle-containing protein marking the eye
DE and the nose) (Dickkopf receptor).
GN KREMEN1 OR KREMEN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Nakamura T., Nakamura T.; to the EMBL/GenBank/DBJ databases.
RL Submitted (JUL-2001)
CC -1- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf
CC to block Wnt/Beta-catenin signaling (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
```

```
CC -1- SIMILARITY: Contains 1 CUB domain.
CC -1- SIMILARITY: Contains 1 Kringle domain.
CC -1- SIMILARITY: Contains 1 WSC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboratio.
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on it
CC use by non-profit institutions as long as its content is in no wa
CC modified and this statement is not removed. Usage by and for commercia
CC entities requires a license agreement (See http://www.isb-sib.ch/announce
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB065090; BAB62003.1; -.
CC GO; GO:0016021; C:integral to membrane; ISS.
CC InterPro; IPR000859; CUB.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR002889; WSC.
CC Pfam; PF00431; CUB; 1.
CC Pfam; PF00051; Kringle; 1.
CC Pfam; PF01822; WSC; 1.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00042; CUB; 1.
CC SMART; SM00130; KR; 1.
CC PROSITE; PS01180; CUB; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS00070; KRINGLE_2; 1.
CC Wnt signaling pathway; Signal; Transmembrane; Kringle.
KW SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 473 KREMEN PROTEIN 1.
FT DOMAIN 21 392 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 393 413 POTENTIAL.
FT DOMAIN 414 473 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 31 114 KRINGLE.
FT DOMAIN 120 210 WSC.
FT DOMAIN 214 321 CUB.
FT CARBOHYD 59 59 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 217 217 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 345 345 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 473 AA; 51869 MW; 9B510857DF856F08 CRC64;
Query Match 20.3%; Score 161; DB 1; Length 473;
Best Local Similarity 45.1%; Pred. No. 1.6e-08;
Matches 32; Conservative 7; Mismatches 28; Indels 4; Gaps 2
QY 47 SKTCYEGNGHYRGKASTDTM--GRPCLPWSATVLOQTYHAHRSALQGLGHKHYCRN 104
DB 29 SPECFTANGADYEGTQSWTALQGGKPCLEWNE--TFQHPYNTLKYPNGEGGLGHEHYCRN 86
QY 105 PDNRRPWCYV 115
DB 87 PDGDSVPWCYV 97
RESULT 24
APOA_MACMU
ID APOA_MACMU STANDARD; PRT; 1420 AA.
AC P14417;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Apolipoprotein(a) (EC 3.4.21.-) (Apo(a)) (Fragment).
GN LPA.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
```



RX MEDLINE=89174660; PubMed=2925643;  
RA Tomlinson J.E., McLean J.W., Iawn R.M.;  
RT "Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of  
RL J. Biol. Chem. 264:5957-5965(1989).  
CC -!- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)  
CC (Lp(a)). It has serine proteinase activity and is able of  
CC autophoretolysis. Inhibits tissue-type plasminogen activator 1.  
CC Lp(a) may be a ligand for megalin/Gp 330.  
CC -!- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and  
CC decorin (By similarity).  
CC -!- PTM: N- and O-glycosylated (By similarity).  
CC -!- DISEASE: Elevated plasma concentrations of apo(a) and its  
CC naturally occurring proteolytic fragments are correlated with  
CC atherosclerosis. Homology with plasminogen kringle IV and V is  
CC thought to underlie the atherogenicity of the protein, because the  
CC fragments are competing with plasminogen for fibrinogen binding.  
CC -!- MISCELLANEOUS: Apo(a) is known to be proteolytically cleaved,  
CC leading to the formation of the so called mini-Lp(a). Apo(a)  
CC fragments accumulate in atherosclerotic lesions, where they may  
CC promote thrombogenesis. O-glycosylation may limit the extent of  
CC proteolytic fragmentation (By similarity).  
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.  
CC -!- SIMILARITY: Contains at least 10 kringle domains.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: J04635; AAA36833.1; -;  
CC PIR: A32869; A32869.  
CC HSSP: P00747; 2PK4.  
CC MEROPS: S01.226;  
CC  
CC InterPro: IPR009003; Cys\_Ser\_trypsin.  
CC InterPro: IPR000001; Kringle.  
CC InterPro: IPR001254; Peptidase\_S1.  
CC InterPro: IPR001314; Peptidase\_S1A.  
CC Pfam: PF00051; Kringle; 1.  
CC Pfam: PF00089; trypsin; 1.  
CC PRINTS: PR00722; CHYMOTRYPSIN.  
CC PRINTS: PR00018; KRINGLE.  
CC ProDom: PD000395; Kringle; 10.  
CC SMART: SM00130; KG; 10.  
CC SMART: SM00020; Tryp\_SPC; 1.  
CC PROSITE: PS00021; KRINGLE 1; 10.  
CC PROSITE: PS00070; KRINGLE 2; 10.  
CC PROSITE: PS0240; TRYPSIN\_DOM; 1.  
CC PROSITE: PS00134; TRYPSIN\_HIS; FALSE\_NEG.  
CC PROSITE: PS00135; TRYPSIN\_SER; FALSE\_NEG.  
CC XW Hydrolase; Serine protease; Lipid transport; Plasma; Glycoprotein;  
CC Kringle; Repeat; Atherosclerosis.  
CC FT NON TER 1 1  
CC FT DOMAIN 49 127 KRINGLE 1.  
CC FT DOMAIN 163 241 KRINGLE 2.  
CC FT DOMAIN 277 355 KRINGLE 3.  
CC FT DOMAIN 391 469 KRINGLE 4.  
CC FT DOMAIN 505 583 KRINGLE 5.  
CC FT DOMAIN 619 697 KRINGLE 6.  
CC FT DOMAIN 725 803 KRINGLE 7.  
CC FT DOMAIN 839 917 KRINGLE 8.  
CC FT DOMAIN 953 1031 KRINGLE 9.  
CC FT DOMAIN 1067 1145 KRINGLE 10.  
CC FT DOMAIN 1191 1420 SERINE PROTEASE.  
CC SQ SEQUENCE 1420 AA; 158367 MW; BE102949E03C5B0E CRC64;  
  
Query Match 20.3%; Score 161; DB 1; Length 1420;  
Best Local Similarity 40.9%; Pred. No. 4 6e-08;  
Matches 36; Conservative 9; Mismatches 31; Indels 12; Gaps 4;

QY 50 CYEGNGHFYRGKASTDTMGSPCLPKNSATVLO--QTYHAHRSDALQLGLGKHYCNRPDN 107  
Db 1068 CYHNGQSYRFTSTVTGTCQSSMTPHQHKRTPENHPNDLLTM-----NYCRNPDA 112  
QY 108 RRRPWCYVQVGLKPLVQE--CMVHDCAD 133  
Db 1123 DTGPMCF--MDPSVRREVCNLTGSD 1147  
  
RESULT 25  
KRM1\_XENLA STANDARD; PRT; 452 AA.  
ID KRM1\_XENLA  
AC Q90Y30;  
DT 10-OCT-2003 (Rel. 42, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Kremen protein 1 precursor (Kringle-containing protein marking the eye  
DE and the nose) (Dickkopf receptor).  
GN KREMEN1.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xenopodinae; Xenopus.  
OC NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Nakamura T.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf  
CC to block Wnt/beta-catenin signaling (By similarity).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).  
CC -!- SIMILARITY: Contains 1 CUB domain.  
CC -!- SIMILARITY: Contains 1 kringle domain.  
CC -!- SIMILARITY: Contains 1 WSC domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collabora  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on it  
CC use by non-profit institutions as long as its content is in no wa  
CC modified and this statement is not removed. Usage by and for commercia  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: AB070851; BAB64294.1; -;  
CC InterPro: IPR000859; CUB.  
CC InterPro: IPR000001; Kringle.  
CC InterPro: IPR002889; WSC.  
CC Pfam: PF00431; CUB; 1.  
CC Pfam: PF00051; kringle; 1.  
CC Pfam: PF01822; WSC; 1.  
CC PRINTS: PR00018; KRINGLE.  
CC ProDom: PD000395; Kringle; 1.  
CC SMART: SM00042; CUB; 1.  
CC SMART: SM00321; KR; 1.  
CC SMART: SM00321; WSC; 1.  
CC PROSITE: PS01180; CUB; 1.  
CC PROSITE: PS00021; KRINGLE 1; 1.  
CC PROSITE: PS00070; KRINGLE 2; 1.  
CC Wnt signaling pathway; Glycoprotein; Kringle;  
CC Transmembrane; Signal.  
CC FT SIGNAL 1 22 POTENTIAL.  
CC FT CHAIN 23 452 KREMEN PROTEIN 1.  
CC FT DOMAIN 23 369 EXTRACELLULAR (POTENTIAL).  
CC FT TRANSMEM 370 390 POTENTIAL.  
CC FT DOMAIN 391 452 CYTOPLASMIC (POTENTIAL).  
CC FT DOMAIN 29 112 KRINGLE.  
CC FT DOMAIN 118 208 WSC.  
CC FT DOMAIN 212 319 CUB.  
CC FT CARBOHYD 43 43 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC FT CARBOHYD 215 215 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC FT CARBOHYD 253 253 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC FT CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).



FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 344 344 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 452 AA; 50188 MW; ED24BCD1AF4564E2 CRC64;

Query Match  
 Best Local Similarity 40.1%; Score 159; DB 1; Length 452;  
 Matches 31; Conservative 9; Mismatches 32; Indels 4; Gaps 2;

QY 42 CEIDSKTCYENGHFYRGKASTDTM--GRPLPWSATVLTQYVHAHRSALGLGLGKH 99  
 DB 22 CDSFHSCEYVNGADYGTQNTQTSLDGKPKLFWNE--TFQHPYNTLYKYPNGEGGLGKH 79

QY 100 NYCRRPDRRRPWCYV 115  
 DB 80 NYCRRPDRGVSPWCYI 95

RESULT 26  
 KRM1\_MOUSE  
 ID KRM1\_MOUSE STANDARD; PRT; 473 AA.  
 AC Q99N43;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Kremen protein 1 precursor (Kringles-containing protein marking the eye and the nose) (Dickkopf receptor).  
 GN KREMEN1 OR KREMEN.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., DEVELOPMENTAL STAGE, AND TISSUE SPECIFICITY.  
 RC TISSUE=Brain, and Kidney;  
 RX MEDLINE=21167372; PubMed=11267660;  
 RA Nakamura T., Aoki S., Kitajima K., Takahashi T., Matsumoto K., Nakamura T.;  
 RT "Molecular cloning and characterization of Kremen, a novel kringles-containing transmembrane protein.";  
 RL Biochim. Biophys. Acta 1518:63-72(2001).  
 CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf to block Wnt/beta-catenin signaling (by similarity).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (potential).  
 CC -!- TISSUE SPECIFICITY: In the adult, widely expressed with high levels in heart, lung, kidney, skeletal muscle and testis.  
 CC -!- DEVELOPMENTAL STAGE: In the embryo, expression is first detected on day 9 and increases up to day 18. Lower levels are found in adult. At 9.5 dpc, expression is localised to the apical ectodermal ridge (AER) of the developing fore- and hindlimb buds, the telencephalon and the first brachial arch. At 10.5 dpc, expression is also observed in the myotome and in sensory tissues such as the nasal pit and optic vesicle.  
 CC -!- SIMILARITY: Contains 1 CUB domain.  
 CC -!- SIMILARITY: Contains 1 kringles domain.  
 CC -!- SIMILARITY: Contains 1 WSC domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AB059617; BAB40968.1;  
 CC HSP; P00747; 1CEA.  
 CC GO; GO:0016021; C:integral to membrane; NAS.  
 CC InterPro; IPR000859; CUB.  
 CC InterPro; IPR002889; WSC.  
 CC Pfam; PF00431; CUB; 1.  
 CC Pfam; PF00051; kringles; 1.

DR Pfam; PF01822; WSC; 1.  
 DR PRINTS; PRO0018; KRINGLE.  
 DR ProDom; PD000395; Kringles; 1.  
 DR SMART; SM00042; CUB; 1.  
 DR SMART; SM00130; KR; 1.  
 DR PROSITE; PS01180; CUB; 1.  
 DR PROSITE; PS00021; KRINGLE\_1; 1.  
 DR PROSITE; PS00070; KRINGLE\_2; 1.  
 KW Wnt signalling pathway; Signal; Transmembrane; Kringles.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT CHAIN 20 473 KREMEN PROTEIN 1.  
 FT DOMAIN 21 392 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 393 413 POTENTIAL.  
 FT DOMAIN 414 473 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 31 114 KRINGLE.  
 FT DOMAIN 120 210 WSC.  
 FT DOMAIN 214 321 CUB.  
 FT CARBOHYD 59 59 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 217 217 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 345 345 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 473 AA; 51716 MW; 586827788BE3FDD1 CRC64;

Query Match  
 Best Local Similarity 20.1%; Score 159; DB 1; Length 473;  
 Matches 31; Conservative 7; Mismatches 26; Indels 4; Gaps 2

QY 50 CYEGNGHFYRGKASTDTM--GRPLPWSATVLTQYVHAHRSALGLGLGKHVCRPDN 107  
 DB 32 CFTANGADYRGTSWTALQGGKPLFWNE--TFQHPYNTLYKYPNGEGGLGKHVCRPDG 89

QY 108 RRRPWCYV 115  
 DB 90 DVSPWCYV 97

RESULT 27  
 ROR1\_HUMAN  
 ID ROR1\_HUMAN STANDARD; PRT; 937 AA.  
 AC Q01973; Q92776;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Tyrosine-protein kinase transmembrane receptor ROR1 precursor (EC 2.7.1.112) (Neurotrophic tyrosine kinase, receptor-related 1).  
 GN ROR1 OR NTRK1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM LONG).  
 RX MEDLINE=93100347; PubMed=1334494;  
 RA Maslakowski P., Carroll R.D.;  
 RT "A novel family of cell surface receptors with tyrosine kinase-like domain.";  
 RL J. Biol. Chem. 267:26181-26190(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM SHORT).  
 RX MEDLINE=97030043; PubMed=8875995;  
 RA Reddy U.R., Phatak S., Pleasure D.;  
 RT "Human neural tissues express a truncated Ror1 receptor tyrosine kinase, lacking both extracellular and transmembrane domains.";  
 RL Oncogene 13:1555-1559(1996).  
 CC -!- FUNCTION: Tyrosine-protein kinase receptor whose role is not yet clear.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=2;

CC Name=Long;  
 CC IsoId=Q01973-1; Sequence=Displayed;  
 CC Name=Short; Synonym=T-ROR1;  
 CC IsoId=Q01973-2; Sequence=VSP\_005008;  
 CC -!- TISSUE SPECIFICITY: Expressed strongly in human heart, lung, and  
 CC kidney, but weakly in the CNS. The short isoform is strongly  
 CC expressed in fetal and adult CNS and in a variety of human  
 CC cancers, including those originating from CNS or PNS  
 CC neuroectoderm.  
 CC -!- DEVELOPMENTAL STAGE: Expressed at high levels during early  
 CC embryonic development. The expression levels drop strongly around  
 CC day 16 and there are only very low levels in adult tissues.  
 CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. ROR  
 CC subfamily.  
 CC -!- SIMILARITY: Contains 1 frizzled (FZ) domain.  
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
 CC -!- SIMILARITY: Contains 1 kringle domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M57675; AAC50275.1; -;  
 CC EMBL; U38894; AAC50714.1; -;  
 CC PIR; A45082; A45082.  
 CC HSP; P00747; 1CEA.  
 CC Genew; HGNC:10256; ROR1.  
 CC MIM; 602336; -;  
 CC GO; GO:0005737; Cytoplasm; TAS.  
 CC GO; GO:0005887; C:integral to plasma membrane; TAS.  
 CC GO; GO:0004714; F:transmembrane receptor protein tyrosine kin. . . ; TAS.  
 CC GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . . ; TAS.  
 CC InterPro; IPR000024; Fz\_domain.  
 CC InterPro; IPR007110; Ig-like.  
 CC InterPro; IPR003598; Ig c2.  
 CC InterPro; IPR000001; Kringle.  
 CC InterPro; IPR007019; Prot\_kinase.  
 CC InterPro; IPR001245; Tyr\_kinase.  
 CC InterPro; IPR008286; Tyr\_pkinase\_AS.  
 CC Pfam; PF01392; Fz; 1.  
 CC Pfam; PF00047; Ig; 1.  
 CC Pfam; PF00051; kringle; 1.  
 CC Pfam; PF00059; pkinase; 1.  
 CC PRINTS; PR00016; KRINGLE.  
 CC PRINTS; PR00109; TYRKINASE.  
 CC ProDom; PD000395; Kringle; 1.  
 CC ProDom; PD000001; Prot\_kinase; 1.  
 CC SMART; SM00408; IGC2; 1.  
 CC SMART; SM00130; KR; 1.  
 CC SMART; SM00219; Tyrc; 1.  
 CC PROSITE; PS00039; Fz; 1.  
 CC PROSITE; PS00835; IG\_LIKE; 1.  
 CC PROSITE; PS00021; KRINGLE\_1; 1.  
 CC PROSITE; PS00070; KRINGLE\_2; 1.  
 CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 CC PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 CC Transferase; Tyrosine-protein kinase; ATP-binding; Receptor;  
 CC Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation;  
 CC Immunoglobulin domain; Alternative splicing.  
 CC SIGNAL 1 29  
 CC CHAIN 30 937  
 CC RECEPTOR ROR1.  
 CC EXTRACELLULAR (POTENTIAL).  
 CC POTENTIAL.  
 CC TRANSMEM 407 427  
 CC DOMAIN 428 937  
 CC CYTOPLASMIC (POTENTIAL).  
 CC DOMAIN 42 147  
 CC IG-LIKE C2-TYPE.  
 CC DOMAIN 165 299  
 CC FZ.  
 CC DOMAIN 312 391  
 CC KRINGLE.

FT DOMAIN 473 746  
 FT SER/THR-RICH. PROTEIN KINASE.  
 FT DOMAIN 753 782  
 FT PRO-RICH. PRO-RICH.  
 FT DOMAIN 784 851  
 FT SER/THR-RICH. PRO-RICH.  
 FT NP\_BIND 479 487  
 FT ATP (BY SIMILARITY).  
 FT BINDING 506 506  
 FT ATP (BY SIMILARITY).  
 FT ACT\_SITE 615 615  
 FT PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 645 645  
 FT DISULFID 79 131  
 FT CARBOHYD 47 47  
 FT CARBOHYD 66 66  
 FT CARBOHYD 184 184  
 FT CARBOHYD 315 315  
 FT CARBOHYD 549 549  
 FT VARSPLIC 1  
 FT Missing (in isoform Short).  
 FT /FTID=VSP\_005008.  
 SQ SEQUENCE 937 AA; 104312 MW; 0D0694DBF23F4773 CRC64;  
 Query Match 20.0%; Score 158.5; DB 1; Length 937;  
 Best Local Similarity 29.7%; Pred. No. 5.4e-08;  
 Matches 43; Conservative 14; Mismatches 51; Indels 37; Gaps 7  
 Qy 2 NELHQP-----SNCCLNGGTCVSNKYFS-----NIHWNC-----PKTGGQH 41  
 Db 236 DETSSVPRDLRCRDECEILENVLCQTEYIFARSNPMLMRLKLPNCEDLPQSPSPAAN 29  
 Qy 42 C-----EIDSKTCYEGNGHFYRGKASTDTMGPCPLPWSATVLOQTYHARS DAL 92  
 Db 296 CIRIGIPMADPINKHKCYNSTGVDTRGTVSVTKSGRCQCPWNS-----QYPHTHTTAL 350  
 Qy 93 QLG--LGKHNRYCRNPDRR--PWCY 114  
 Db 351 RFPELNGHGYCRNPGNQKEAPWC 375  
 RESULT 28  
 ID ROR1\_MOUSE STANDARD; PRT; 937 AA.  
 AC Q9Z139;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Tyrosine-protein kinase transmembrane receptor ROR1 precursor  
 DE (EC 2.7.1.112) (Neurotrophic tyrosine kinase, receptor-related 1)  
 DE (mROR1).  
 GS ROR1 OR NTRK1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99248426; PubMed=10231392;  
 RA Oishi I., Takeuchi S., Hashimoto R., Nagabukuro A., Ueda T., Liu Z.J.,  
 RA Hata T., Akira S., Matsuda Y., Yamamura H., Otani H., Minami Y.,  
 RT "Spatio-temporally regulated expression of receptor tyrosine kinases,  
 RT mRor1, mRor2, during mouse development: implications in development  
 RT and function of the nervous system."  
 RL Genes Cells 4:41-56(1999).  
 CC -!- FUNCTION: Tyrosine-protein kinase receptor whose role is not yet  
 CC clear.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. ROR  
 CC subfamily.  
 CC -!- SIMILARITY: Contains 1 frizzled (FZ) domain.  
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
 CC -!- SIMILARITY: Contains 1 kringle domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboratio  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on it  
 CC use by non-profit institutions as long as its content is in no wa







```

DR InterPro: IPR000003; Cys ser trypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003609; Pan_app.
DR InterPro: IPR001254; Peptidase_S1.
DR InterPro: IPR001314; Peptidase_S1A.
DR InterPro: IPR003966; Peptidase_S1A_pr.
DR Pfam: PF00051; kringle; 5.
DR Pfam: PF00024; PAN; 1.
DR Pfam: PF00089; tyrosin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR PRINTS: PR01505; PROTHROMBIN.
DR ProDom: PD000395; Kringle; 5.
DR SMART: SM00130; KR; 5.
DR SMART: SM00473; PAN AP; 1.
DR SMART: SM00020; TVD_SPC; 1.
DR PROSITE: PS00021; KRINGLE_1; 5.
DR PROSITE: PS00070; KRINGLE_2; 5.
DR PROSITE: PS0240; TRYPSIN_DOM; 1.
DR PROSITE: PS0134; TRYPSIN_HIS; 1.
DR PROSITE: PS0135; TRYPSIN_SER; 1.
DR KMW: Tissue remodeling; Blood coagulation; Plasma; Glycoprotein; Fibrinolysis;
DR KMW: Tissue remodeling; Kringle; Zymogen; Repeat.
FT CHAIN 1 560 PLASMIN LIGHT CHAIN B.
FT CHAIN 561 790 PLASMIN HEAVY CHAIN A.
FT DOMAIN 561 790 SERINE PROTEASE.
FT DOMAIN 84 162 KRINGLE 1.
FT DOMAIN 166 243 KRINGLE 2.
FT DOMAIN 256 333 KRINGLE 3.
FT DOMAIN 358 435 KRINGLE 4.
FT DOMAIN 461 540 KRINGLE 5.
FT ACT_SITE 602 602 CHARGE RELAY SYSTEM.
FT ACT_SITE 645 645 CHARGE RELAY SYSTEM.
FT ACT_SITE 740 740 CHARGE RELAY SYSTEM.
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .).
FT CARBOHYD 340 340 /FTID-CAR 000019.
FT CARBOHYD 340 340 O-LINKED (GALNAC. . .).
FT SEQUENCE 790 AA; 88592 MW; F04EA0574BCD58E CRC64;

Query Match 19.5%; Score 155; DB 1; Length 790;
Best Local Similarity 27.2%; Pred. No. 1e-07;
Matches 44; Conservative 16; Mismatches 56; Indels 46; Gaps 9;

QY 5 HQVPSNCDCLNGTCVSNKYFSN-----IHWK-NCPKKGQHCETDK----- 46
Db 289 NRTPENPCKN----LEENYCRNPDGETAPWCTTDSVVRWDYCKIPSCGSSTTSTEHL 344
QY 47 -----SKTCYEGNGHFYRGKASTDTMGPCPLPWN SATVLQTYHAHR----SDALQ 93
Db 345 APVPEPTPAQDCYRGNGESYRGTSSTITGRKQSW-----VSMTPHREKTPGNFPN 399
QY 94 LGLOKHNYCRNPNRRPWCYVQGLKPLV--QECMVHDCAD 133
Db 400 AGL-TMNYCRNPADKSPWCYT---TDPVRVWEYCNLKKSG 437

RESULT 32
PLMN_HUMAN
ID PLMN_HUMAN STANDARD; PRT; 810 AA.
AC P00747;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].
GN PLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANT ASN-472.

```

```

RX MEDLINE=90202879; PubMed=2318848;
RA Petersen T.E., Martzen M.R., Ichinose A., Davie E.W.;
RT "Characterization of the gene for human plasminogen, a key proenzyme
in the fibrinolytic system.";
RL J. Biol. Chem. 265:6104-6111(1990).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=87162490; PubMed=3030813;
RA Forsgren M., Raden B., Israelsson M., Larsson K., Heden L.-O.;
RT "Molecular cloning and characterization of a full-length cDNA clone
for human plasminogen.";
RL FEBS Lett. 213:254-260(1987).
RN [3]
RN SEQUENCE FROM N.A., AND VARIANTS LYS-57; GLN-133; HIS-261; TRP-408;
RP ASN-472; VAL-494 AND TRP-523.
RA Rieder M.J., Armel T.Z., Carrington D.P., Ozuna M., Kuidanek S.A.,
RA Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 20-810, AND VARIANT ASN-472.
RA Sottrup-Jensen L., Petersen T.E., Magnusson S.;
RL Submitted (JUL-1977) to the PIR data bank.
RN [5]
RP SEQUENCE OF 292-810 FROM N.A.
RX MEDLINE=85023311; PubMed=6148961;
RA Malinowski D.P., Sadler J.E., Davie E.W.;
RT "Characterization of a complementary deoxyribonucleic acid coding for
human and bovine plasminogen.";
RL Biochemistry 23:4243-4250(1984).
RN [6]
RP SEQUENCE OF 20-100.
RX MEDLINE=75093329; PubMed=122932;
RA Wiman B., Wallen P.;
RT "Structural relationship between 'glutamic acid' and 'lysine' forms
of human plasminogen and their interaction with the NH2-terminal
activation peptide as studied by affinity chromatography.";
RL Eur. J. Biochem. 50:489-494(1975).
RN [7]
RP SEQUENCE OF 95-580; 581-626; 657-700 AND 732-810, AND VARIANT ASN-472.
RA Sottrup-Jensen L., Claessens H., Zaidel M., Petersen T.E., Magnusson S.;
RL (in) Davidson J.F., Rowan R.M., Samama M.M., Desnoyers P.C. (eds.);
Progress in Chemical Fibrinolysis and Thrombolysis, pp.3:191-209,
RL Raven Press, New York (1978).
RN [8]
RP SEQUENCE OF 483-604.
RX MEDLINE=76043692; PubMed=126863;
RA Wiman B., Wallen P.;
RT "Amino-acid sequence of the cyanogen-bromide fragment from human
plasminogen that forms the linkage between the plasmin chains.";
RL Eur. J. Biochem. 58:539-547(1975).
RN [9]
RP SEQUENCE OF 581-810.
RX MEDLINE=77252545; PubMed=142009;
RA Wiman B.;
RT "Primary structure of the B-chain of human plasmin.";
RL Eur. J. Biochem. 76:129-137(1977).
RN [10]
RP ACTIVE SITE.
RX MEDLINE=73149248; PubMed=4694729;
RA Robbins K.C., Bernabe P., Arzadon L., Summari L.;
RT "The primary structure of human plasminogen. II. The histidine loop
of human plasmin: light (B) chain active center histidine sequence.";
RL J. Biol. Chem. 248:1631-1633(1973).
RN [11]
RP ACTIVE SITE.
RX MEDLINE=69234739; PubMed=4240117;
RA Groskopf W.R., Summari L., Robbins K.C.;
RT "Studies on the active center of human plasmin. Partial amino acid
sequence of a peptide containing the active center serine residue.";
RL J. Biol. Chem. 244:3590-3597(1969).
RN [12]
RP OMEGA-AMINOCARBOXYLIC ACID-BINDING SITES.
RX MEDLINE=82213905; PubMed=6919539;

```

RA Trexler M., Vali Z., Patthy L.;  
RT "Structure of the omega-aminocarboxylic acid-binding sites of human  
RT plasminogen. Arginine 70 and aspartic acid 56 are essential for  
RT binding of ligand by kringle 4.";  
RL J. Biol. Chem. 257:7401-7406(1982).  
RN [13]  
RP FIBRIN AND OMEGA-AMINOCARBOXYLIC ACID BINDING SITES.  
RX MEDLINE=85054794; PubMed=6094526;  
RA Vali Z., Patthy L.;  
RT "The fibrin-binding site of human plasminogen. Arginines 32 and 34  
RT are essential for fibrin affinity of the kringle 1 domain.";  
RL J. Biol. Chem. 259:13690-13694(1984).  
RN [14]  
RP PHOSPHORYLATION SITE SER-597.  
RX MEDLINE=97345939; PubMed=9201958;  
RA Wang H., Prorok M., Brethauer R.K., Castellino F.J.;  
RT "Serine-578 is a major phosphorylation locus in human plasma  
RT plasminogen.";  
RL Biochemistry 36:8100-8106(1997).  
RN [15]  
RP CARBOHYDRATE-LINKAGE SITES.  
RX MEDLINE=88185329; PubMed=3356193;  
RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,  
RA Gerwig G.J., van Halbeek H., Vliegenthart J.F.;  
RT "The N- and O-linked carbohydrate chains of human, bovine and porcine  
RT plasminogen. Species specificity in relation to sialylation and  
RT fucosylation patterns.";  
RL Eur. J. Biochem. 173:57-63(1988).  
RN [16]  
RP CARBOHYDRATE-LINKAGE SITE SER-268.  
RX MEDLINE=97207306; PubMed=9034441;  
RA Pirie-Shepherd S.R., Stevens R.D., Andon N.L., Enghild J.J.,  
RA Pizzo S.V.;  
RT "Evidence for a novel O-linked sialylated trisaccharide on Ser-248 of  
RT human plasminogen 2.";  
RL J. Biol. Chem. 272:7408-7411(1997).  
RN [17]  
RP CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.  
RX MEDLINE=95042728; PubMed=7525077;  
RA O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,  
RA Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;  
RT "Angiostatin: a novel angiogenesis inhibitor that mediates the  
RT suppression of metastases by a Lewis lung carcinoma.";  
RL Cell 79:315-328(1994).  
RN [18]  
RP CHARACTERIZATION OF ANGIOSTATIN.  
RX MEDLINE=97238710; PubMed=1602221;  
RA Sim B.K., O'Reilly M.S., Liang H., Fortier A.H., He W., Madsen J.W.,  
RA Lapcevic R., Nacy C.A., Liang H., Fortier A.H., He W., Madsen J.W.,  
RT "A recombinant human angiostatin protein inhibits experimental primary  
RT and metastatic cancer.";  
RL Cancer Res. 57:1329-1334(1997).  
RN [19]  
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 374-461.  
RX MEDLINE=92031502; PubMed=1657149;  
RA Mulichak A.M., Tulinsky A., Ravichandran K.G.;  
RT "Crystal and molecular structure of human plasminogen kringle 4  
RT refined at 1.9-A resolution.";  
RL Biochemistry 30:10576-10588(1991).  
RN [20]  
RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 374-461.  
RX MEDLINE=92031503; PubMed=1657149;  
RA Wu T.-P., Padmanabhan K., Tulinsky A., Mulichak A.M.;  
RT "The refined structure of the epsilon-aminocaproic acid complex of  
RT human plasminogen kringle 4.";  
RL Biochemistry 30:10589-10594(1991).  
RN [21]  
RP X-RAY CRYSTALLOGRAPHY (1.67 ANGSTROMS) OF 376-454.  
RA Stec B., Yamano A., Whitlow M., Teeter M.M.;  
RT "Structure of human plasminogen kringle 4 at 1.68 Angstrom and 277 K.  
RT A possible structural role of disordered residues.";  
RL Acta Crystallogr. D 53:169-178(1997).  
RN [22]

RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 102-181.  
RX MEDLINE=96180681; PubMed=8611560;  
RA Mathews I.I., Vanderhoff-Hanaver P., Castellino F.J., Tulinsky A.;  
RT "Crystal structures of the recombinant kringle 1 domain of human  
RT plasminogen in complexes with the ligands epsilon-aminocaproic acid  
RT and trans-4-(aminomethyl)cyclohexane-1-carboxylic Acid.";  
RL Biochemistry 35:2567-2576(1996).  
RN [23]  
RP X-RAY CRYSTALLOGRAPHY (1.66 ANGSTROMS) OF 480-563.  
RX MEDLINE=98198034; PubMed=9521645;  
RA Chang Y., Mochalkin I., McCance S.G., Cheng B., Tulinsky A.,  
RA Castellino F.J.;  
RT "Structure and ligand binding determinants of the recombinant kringle  
RT 5 domain of human plasminogen.";  
RL Biochemistry 37:3258-3271(1998).  
RN [24]  
RP STRUCTURE BY NMR OF 96-184.  
RX MEDLINE=94237157; PubMed=8181475;  
RA Rejante M.R., Llinas M.;  
RT "1H-NMR assignments and secondary structure of human plasminogen  
RT kringle 1.";  
RL Eur. J. Biochem. 221:927-937(1994).  
RN [25]  
RP STRUCTURE BY NMR OF 96-184.  
RX MEDLINE=94237158; PubMed=8181476;  
RA Rejante M.R., Llinas M.;  
RT "Solution structure of the epsilon-aminohexanoic acid complex of  
RT human plasminogen kringle 1.";  
RL Eur. J. Biochem. 221:939-949(1994).  
RN [26]  
RP STRUCTURE BY NMR OF 183-354.  
RX MEDLINE=96194156; PubMed=8652577;  
RA Soehndel S., Hu C.-K., Marti D., Affolter M., Schaller J., Llinas M.,  
RA Rickli E.E.;  
RT "Recombinant gene expression and 1H NMR characteristics of the  
RT kringle (2 + 3) supermodule: spectroscopic/functional individuality  
RT of plasminogen kringle domains.";  
RL Biochemistry 35:2357-2364(1996).  
RN [27]  
RP STRUCTURE BY NMR OF 374-461.  
RX MEDLINE=90219023; PubMed=2157850;  
Query Match 19.48; Score 154; DB 1; Length 810;  
Best Local Similarity 28.08; Pred. No. 1.3e-07; Indels 46; Gaps 10  
Matches 45; Conservative 20; Mismatches 50;  
QY 5 HQVPSNCDCLNGGTCVSNKYFSN-----IHWNCNP-KKFGQHCEI---DKS-----47  
DB 308 NRTFENPCKN---LDENYCRNPDGKRAPWCHTNSQVWEYCKIPSCDSSPVSTBQLA 363  
QY 48 -----KTCYEGNGHFYRGKASTDTMGRPCLPFNSTATVLTQTYHAHR---SDALQ 93  
DB 364 PTAPPELTVPVQDCYHGDGSGYRGTSTTTTGKKCQSWSS-----MTPHRHQKTPENYFN 416  
QY 94 LGLGKHNYCRNPNRPPWCVVQVGLKPLV--QECMVHDC A 132  
DB 419 AGL-TMNYCRNPDADKGPWCFT---TDSVWEYCNLKKCS 455  
RESULT 33  
PLAN ERIEU  
ID PLAN ERIEU STANDARD; PRT; 810 AA.  
AC Q29485;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Plasminogen precursor (EC 3.4.21.7).  
GN PLG.  
OS Erinaceus europaeus (Western European hedgehog).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Insectivora; Erinaceidae; Erinaceus.  
OX NCBI\_TaxID=9365;  
RN [1]



SEQUENCE FROM N.A.

TISSUE=Liver;  
MEDLINE=96025778; PubMed=7592597;  
Lawn R.M., Boonmark N.W., Schwartz K., Lindahl G.E., Wade D.P.,  
Byrne C.D., Fong K.J., Meer K., Patchy L.;  
"The recurring evolution of lipoprotein(a): Insights from cloning of  
hedgehog apolipoprotein(a)".  
J. Biol. Chem. 270:24004-24009(1995).  
[2]

REVISONS.

Lawn R.M.;  
Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.

-!- FUNCTION: plasmin dissolves the fibrin of blood clots and acts as  
a proteolytic factor in a variety of other processes including  
embryonic development, tissue remodeling, tumor invasion, and  
inflammation; in ovulation it weakens the walls of the Graafian  
follicle. It activates the urokinase-type plasminogen activator,  
collagenases and several complement zymogens, such as C1 and C5.  
It cleaves fibrin, fibronectin, thrombospondin, laminin and von  
Willebrand factor.

-!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;  
higher selectivity than trypsin. Converts fibrin into soluble  
products.

-!- ENZYME REGULATION: Converted into plasmin by plasminogen  
activators, both plasminogen and its activator being bound to  
fibrin. Cannot be activated with streptokinase.

-!- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin  
immediately after dissociation from the clot.

-!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.

-!- SIMILARITY: Contains 5 kringle domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; U33171; AAC48717.1; -  
PIR; I46260; I46260.  
HSSP; P00747; IPMK.

MEROPS; S01.233; -  
InterPro; IPR009003; Cys\_Ser\_trypsin.  
InterPro; IPR000001; Kringle.  
InterPro; IPR003014; PAN.  
InterPro; IPR003609; Pan\_app.  
InterPro; IPR001254; Peptidase\_S1.  
InterPro; IPR001314; Peptidase\_S1A.  
InterPro; IPR003966; Peptidase\_S1A\_pr.  
Pfam; PF00051; Kringle; 5.  
Pfam; PF00024; PAN; 1.  
Pfam; PF00089; trypsin; 1.  
PRINTS; PR00722; CHYMOTRYPSIN.  
PRINTS; PR00018; KRINGLE  
PRINTS; PR01505; PROTHROMBIN.  
ProDom; PD000395; Kringle; 5.  
SMART; SM00130; KR; 5.  
SMART; SM00473; PAN AP; 1.  
SMART; SM00020; Tryp\_SPC; 1.  
PROSITE; PS00021; KRINGLE\_1; 5.  
PROSITE; PS00070; KRINGLE\_2; 5.  
PROSITE; PS00240; TRYPSIN\_DOM; 1.  
PROSITE; PS00134; TRYPSIN\_HIS; 1.  
PROSITE; PS00135; TRYPSIN\_SER; 1.  
Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;  
Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;  
Signal.

1 19 BY SIMILARITY.  
FT CHAIN 20 810 PLASMINOGEN.  
FT CHAIN 20 582 PLASMIN HEAVY CHAIN A (BY SIMILARITY).  
FT CHAIN 583 810 PLASMIN LIGHT CHAIN B (BY SIMILARITY).  
FT DOMAIN 583 810 SERINE PROTEASE.



between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; J04697; AAA36901.1; -  
PIR; B32869; B30848.  
HSP; P00747; 1PMK.  
MEROPS; S01.233; -  
InterPro; IPR009003; Cys Ser trypsin.  
InterPro; IPR000001; Kringle.  
InterPro; IPR003014; PAN.  
InterPro; IPR003609; PAN app.  
InterPro; IPR001254; Peptidase S1.  
InterPro; IPR001314; Peptidase S1A.  
InterPro; IPR003966; Peptidase\_S1A\_pr.  
Pfam; PF00051; kringle; 5.  
Pfam; PF00024; PAN; 1.  
Pfam; PF00089; trypsin; 1.  
PRINTS; P00722; CHYMOTRYPSIN.  
PRINTS; P00018; KRINGLE.  
PRINTS; P01505; PROTHROMBIN.  
ProDom; PD000395; Kringle; 5.  
SMART; SM00130; KR; 4.  
SMART; SM00473; PAN AP; 1.  
SMART; SM00020; Tryp\_Spc; 1.  
PROSITE; PS00021; KRINGLE\_1; 5.  
PROSITE; PS00070; KRINGLE\_2; 5.  
PROSITE; PS00240; TRYPSIN\_DOM; 1.  
PROSITE; PS00134; TRYPSIN\_HIS; 1.  
PROSITE; PS00135; TRYPSIN\_SER; 1.  
Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;  
Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;  
Signal.  
FT SIGNAL 1 19  
FT CHAIN 20 810 PLASMINOGEN.  
FT CHAIN 20 580 PLASMIN HEAVY CHAIN A.  
FT PEPTIDE 20 96 ACTIVATION PEPTIDE.  
FT CHAIN 97 580 PLASMIN SHORT FORM OF CHAIN A.  
FT CHAIN 581 810 PLASMIN LIGHT CHAIN B.  
FT DOMAIN 103 181 KRINGLE 1.  
FT DOMAIN 184 262 KRINGLE 2.  
FT DOMAIN 275 352 KRINGLE 3.  
FT DOMAIN 377 454 KRINGLE 4.  
FT DOMAIN 481 560 KRINGLE 5.  
FT DOMAIN 581 810 SERINE PROTEASE.  
FT ACT\_SITE 622 622 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 665 665 CHARGE RELAY SYSTEM.  
FT BINDING 136 136 OMEGA-AMINOCARBOXYLIC ACIDS.  
FT BINDING 158 158 OMEGA-AMINOCARBOXYLIC ACIDS.  
FT BINDING 172 172 OMEGA-AMINOCARBOXYLIC ACIDS.  
FT BINDING 432 432 OMEGA-AMINOCARBOXYLIC ACIDS.  
FT BINDING 445 445 OMEGA-AMINOCARBOXYLIC ACIDS.  
FT BINDING 134 134 FIBRIN.  
FT BINDING 136 136 FIBRIN.  
FT DISULFID 49 73 BY SIMILARITY.  
FT DISULFID 53 61 BY SIMILARITY.  
FT DISULFID 103 181 BY SIMILARITY.  
FT DISULFID 124 164 BY SIMILARITY.  
FT DISULFID 152 176 BY SIMILARITY.  
FT DISULFID 185 262 BY SIMILARITY.  
FT DISULFID 188 316 BY SIMILARITY.  
FT DISULFID 206 245 BY SIMILARITY.  
FT DISULFID 234 257 BY SIMILARITY.  
FT DISULFID 275 352 BY SIMILARITY.  
FT DISULFID 296 335 BY SIMILARITY.  
FT DISULFID 324 347 BY SIMILARITY.  
FT DISULFID 377 454 BY SIMILARITY.  
FT DISULFID 398 437 BY SIMILARITY.  
FT DISULFID 426 449 BY SIMILARITY.

FT DISULFID 481 560 BY SIMILARITY.  
FT DISULFID 502 543 BY SIMILARITY.  
FT DISULFID 531 555 BY SIMILARITY.  
FT DISULFID 567 685 BY SIMILARITY.  
FT DISULFID 577 583 BY SIMILARITY.  
FT DISULFID 607 623 BY SIMILARITY.  
FT DISULFID 699 766 BY SIMILARITY.  
FT DISULFID 729 745 BY SIMILARITY.  
FT DISULFID 756 784 BY SIMILARITY.  
FT CARBOHYD 365 365 O-LINKED (GALNAC. .) (BY SIMILARITY).  
SQ SEQUENCE 810 AA; 90255 MW; A75ELCS1A1AOP24A CRC64;  
Query Match 19.2%; Score 152.5; DB 1; Length 810;  
Best Local Similarity 31.9%; Pred. No. 1.se-07;  
Matches 45; Conservative 10; Mismatches 61; Indels 25; Gaps 6  
QY 10 NCDLNGGTCVSNKYPNSIHWNCNPKKFGGQ-----HCEIDKSKTCYEGNGH 56  
DB 428 NPDADKGPWCFTTDPVSRWEYCNLKKCSGTESVAAPPVVAQLPDAETSEEDCMFGNCK 487  
QY 57 FYRGKASTDTMGPRCLPWSATVLOQTYAHR-----SDALQLGLGKHNCRPD-NRRPP 111  
DB 488 GYRGKATTTVTGTCQEWAA-----QEPHSHRIFTETNPRAGLEK-NYCRNPDGVDGVP 541  
QY 112 WCYVQVGLKPLVQECMVHDC A 132  
DB 542 WCYT-TNPKLFDYCDVPCA 561  
RESULT 35  
PLMN RAT STANDARD; PRT; 169 AA.  
ID PLMN RAT AC Q0117;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Plasminogen (EC 3.4.21.7) (Fragment).  
GN PLG.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
NCBI\_TaxID=10116;  
[1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=91250378; PubMed=1645711;  
RA Kanalas J.J., Makker S.P.;  
RT "Identification of the rat Heymann nephritis autoantigen (GP330) as a receptor site for plasminogen.";  
RL J. Biol. Chem. 266:10825-10829(1991).  
CC -!- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as a proteolytic factor in a variety of other processes including embryonic development, tissue remodeling, tumor invasion, and inflammation; in ovulation it weakens the walls of the Graafian follicle. It activates the urokinase-type plasminogen activator, collagenases and several complement zymogens, such as C1 and C5. It cleaves fibrin, fibronectin, thrombospondin, laminin and von Willebrand factor.  
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa; higher selectivity than trypsin. Converts fibrin into soluble products.  
CC -!- ENZYME REGULATION: Converted into plasmin by plasminogen activators, both plasminogen and its activator being bound to fibrin. Cannot be activated with streptokinase.  
CC -!- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin immediately after dissociation from the clot.  
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no wa



```
DR PRINTS; PRO0018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00321; WSC; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS00021; KRINGLE; 1; 1.
DR PROSITE; PS00070; KRINGLE; 2; 1.
KW Wnt signaling pathway; Glycoprotein; Signal; Transmembrane;
KW Alternative splicing.
FT SIGNAL 1 25
FT CHAIN 26 462
FT DOMAIN 26 364
FT TRANSMEM 365 387
FT DOMAIN 388 462
FT DOMAIN 35 119
FT DOMAIN 121 215
FT DOMAIN 219 326
FT CARBOHYD 49 49
FT CARBOHYD 222 222
FT CARBOHYD 244 244
FT CARBOHYD 351 351
FT VARSPLIC 394 424
FT FTID-VSP 050509.
FT Missing (in isoform 2).
FT FTID-VSP 050510.
FT ARVSTVIAVSVLLLLGLRLPRLRRSLIAPKGGPPALG
FT ASRGPRSWAWVY -> GAVCWLRKGRPRRLGLPCAPGEAG
FT LGGTSPEGWPCAPPPTPTLRVLPRATGL (in
FT isoform 3).
FT FTID-VSP 050511.
FT Missing (in isoform 3).
FT FTID-VSP 050512.
FT ARVSTVIAVSVLLLLGLRLPRLRRSLIAP -> GEAG
FT ADGSGSRPLAPILTAACVCPQGSRR (in isoform
FT 4).
FT FTID-VSP 050513.
FT Missing (in isoform 4).
FT FTID-VSP 050514.
FT Missing (in Ref. 2; BAC11365).
FT A -> D (in Ref. 2; BAC11365).
FT SEQUENCE 462 AA; 4894 MW; CE33015917A9AA68 CRC64;
Query Match 18.7%; Score 148.5; DB 1; Length 462;
Best Local Similarity 42.0%; Pred. No. 2.6e-07;
Matches 29; Conservative 9; Mismatches 26; Indels 5; Gaps 2;
QY 50 CYEGNGHYRG---KASDTMGRCLPWSATVLOQTTHARSALQLGLGKHYCRNPD 106
Db 36 CFQVNGADYRGHQNRGPRGAGRCFLFDQTO--QHSYSASDPHGRLGHAHFCRNPD 93
QY 107 NRRFWCYV 115
Db 94 GDVQPCYV 102
RESULT 37
ROR2_HUMAN
ID ROR2_HUMAN STANDARD; PRT; 943 AA.
AC Q01974; Q9HAY7; Q9HB61;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Tyrosine-protein kinase transmembrane receptor ROR2 precursor
DE (EC 2.7.1.112) (Neurotrophic tyrosine kinase, receptor-related 2).
GN ROR2 OR NTRK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RX MEDLINE-93100347; PubMed-1334494;
RA Masikowski P., Carroll R.D.;
RT "A novel family of cell surface receptors with tyrosine kinase-like
RL domain.", Chem. 267:26181-26190 (1992).
RN [2]
RP SEQUENCE OF 34-943 FROM N.A., AND VARIANT ILE-819.
RX MEDLINE-20164326; PubMed-10700182;
RA Oldridge M., Fortuna A.M., Maringa M., Propping P., Mansour S.,
RA Pollitt C., DeChiara T.M., Kimble R.B., Valenzuela D.M.,
RA Yancopoulos G.D., Wilkie A.O.M.;
RT "Dominant mutations in ROR2, encoding an orphan receptor tyrosine
RL kinase, cause brachydactyly type B.",
RN Nat. Genet. 24:275-278 (2000).
RN [3]
RP SEQUENCE OF 34-574 FROM N.A., AND VARIANT THR-245.
RX MEDLINE-20442029; PubMed-10986040;
RA Schwabe G.C., Tinschert S., Buschow C., Meinecke P., Wolff G.,
RA Gillissen-Kaesbach G., Oldridge M., Wilkie A.O.M., Koemec R.,
RA Mundlos S.;
RT "Distinct mutations in the receptor tyrosine kinase gene ROR2 cause
RL brachydactyly type B.",
RN Am. J. Hum. Genet. 67:822-831 (2000).
RN [4]
RP VARIANTS RRS CYS-184; TRP-189; TRP-366 AND LYS-620.
RX MEDLINE-20392394; PubMed-10932186;
RA Arzal A.R., Rajab A., Fenske C.D., Oldridge M., Elanko N.,
RA Ternes-Pereira E., Tveysuez B., Munday V.A., Patton M.A.,
RA Wilkie A.O.M., Jeffery S.;
RT "Recessive Robinow syndrome, allelic to dominant brachydactyly type B,
RL is caused by mutation of ROR2.",
RN Nat. Genet. 25:419-422 (2000).
RN [5]
RP VARIANT RRS TYR-182.
RX MEDLINE-20392395; PubMed-10932187;
RA van Bokhoven H., Celli J., Kayserili H., van Beusekom E., Balci S.,
RA Brussel W., Skovby F., Kerr B., Percin E.F., Akarsu N., Brunner H.G.;
RT "Mutation of the gene encoding the ROR2 tyrosine kinase causes
RL autosomal recessive Robinow syndrome.",
RN Nat. Genet. 25:423-426 (2000).
RN [6]
RP ERRATUM.
RA van Bokhoven H., Celli J., Kayserili H., van Beusekom E., Balci S.,
RA Brussel W., Skovby F., Kerr B., Percin E.F., Akarsu N., Brunner H.G.;
RN Nat. Genet. 26:383-383 (2000).
CC -I- FUNCTION: Tyrosine-protein kinase receptor which may be involved
CC in the early formation of the chondrocytes. It seems to be
CC required for cartilage and growth plate development.
CC -I- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -I- SUBCELLULAR LOCATION: Type I membrane protein.
CC -I- DEVELOPMENTAL STAGE: Expressed at high levels during early
CC embryonic development. The expression levels drop strongly around
CC day 16 and there are only very low levels in adult tissues.
CC -I- DISEASE: Defects in ROR2 are a cause of brachydactyly type B1
CC (BDB1) [MIM:113000]. BDB1 is an autosomal dominant skeletal
CC disorder characterized by hypoplasia/aplasia of distal phalanges
CC and nails. In BDB1 the middle phalanges are short but in addition
CC the terminal phalanges are rudimentary or absent. Both fingers and
CC toes are affected. The thumbs and big toes are usually deformed.
CC -I- DISEASE: Defects in ROR2 are a cause of recessive Robinow syndrome
CC (RRS) [MIM:269310]. RRS is an autosomal disorder characterized by
CC skeletal dysplasia with generalized limb bone shortening,
CC segmental defects of the spine, brachydactyly and a dysmorphic
CC facial appearance.
CC -I- SIMILARITY: Belongs to the Tyr family of protein kinases. ROR
CC subfamily.
CC -I- SIMILARITY: Contains 1 frizzled (FZ) domain.
CC -I- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -I- SIMILARITY: Contains 1 kringle domain.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaborative
```

between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

ENBL; M97639; AAA60276.1; --	
ENBL; AF294796; AAG01184.2; --	
ENBL; AF254747; AAG01184.2; JOINED.	
ENBL; AF254748; AAG01184.2; JOINED.	
ENBL; AF254749; AAG01184.2; JOINED.	
ENBL; AF254750; AAG01184.2; JOINED.	
ENBL; AF254751; AAG01184.2; JOINED.	
ENBL; AF254752; AAG01184.2; JOINED.	
ENBL; AF254753; AAG01184.2; JOINED.	
ENBL; AF279762; AAG33132.1; --	
ENBL; AF279755; AAG33132.1; JOINED.	
ENBL; AF279756; AAG33132.1; JOINED.	
ENBL; AF279757; AAG33132.1; JOINED.	
ENBL; AF279758; AAG33132.1; JOINED.	
ENBL; AF279759; AAG33132.1; JOINED.	
ENBL; AF279760; AAG33132.1; JOINED.	
ENBL; AF279761; AAG33132.1; JOINED.	
PIR; B45082; B45082.	
HSSP; P00747; ICRN.	
Genew; HGNC.10257; ROR2.	
MIM; 602337; --	
MIM; 113000; --	
MIM; 268310; --	
GO; GO:0005987; C:integral to plasma membrane; TAS.	
GO; GO:0004714; P:transmembrane receptor protein tyrosine kin. . . ; TAS.	
GO; GO:0007275; P:development; TAS.	
GO; GO:0007165; P:signal transduction; TAS.	
InterPro; IPR000024; Fz_domain.	
InterPro; IPR007110; Ig-like.	
InterPro; IPR003598; Ig_c2.	
InterPro; IPR000001; Kringle.	
InterPro; IPR000719; Prot_kinase.	
InterPro; IPR001245; Tyr_kinase.	
InterPro; IPR008266; Tyr_kinase_AS.	
Pfam; PF01392; Fz; 1.	
Pfam; PF00047; Ig; 1.	
Pfam; PF00051; kringle; 1.	
Pfam; PF00069; pkinase; 1.	
PRINTS; PR00018; KRINGLE.	
PRINTS; PR00109; TYRKINASE.	
ProDom; PD000395; kringle; 1.	
ProDom; PD000001; Prot_kinase; 1.	
SMART; SM00408; IGC2; 1.	
SMART; SM00330; KR; 1.	
SMART; SM00219; TyKc; 1.	
PROSITE; PS50038; Fz; 1.	
PROSITE; PS50835; IG LIKE; 1.	
PROSITE; PS50021; KRINGLE_1; 1.	
PROSITE; PS50070; KRINGLE_2; 1.	
PROSITE; PS00107; PROTEIN KINASE ATP; FALSE NEG.	
PROSITE; PS50011; PROTEIN KINASE DOM; 1.	
PROSITE; PS00109; PROTEIN KINASE TYR; 1.	
TransEnzyme; Tyrosine-protein kinase; ATP-binding; Receptor;	
Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation;	
Immunoglobulin domain; Developmental protein; Polymorphism;	
Disease mutation.	
FT SIGNAL	33
FT CHAIN	34 943
FT FT	
DOMAIN	34 403
TRANSNEM	404 424
FT DOMAIN	425 943
FT DOMAIN	55 145
FT DOMAIN	169 303
FT DOMAIN	316 394
FT DOMAIN	473 746
FT FT	
DOMAIN	403 424
TRANSNEM	425 943
FT DOMAIN	55 145
FT DOMAIN	169 303
FT DOMAIN	316 394
FT DOMAIN	473 746
FT FT	
DOMAIN	403 424
TRANSNEM	425 943
FT DOMAIN	55 145
FT DOMAIN	169 303
FT DOMAIN	316 394
FT DOMAIN	473 746
FT FT	
DOMAIN	403 424
TRANSNEM	425 943
FT DOMAIN	55 145
FT DOMAIN	169 303
FT DOMAIN	316 394
FT DOMAIN	473 746
FT FT	
DOMAIN	403 424
TRANSNEM	425 943
FT DOMAIN	55 145
FT DOMAIN	169 303
FT DOMAIN	316 394
FT DOMAIN	473 746
FT FT	
DOMAIN	403 424
TRANSNEM	425 943
FT DOMAIN	55 145
FT DOMAIN	169 303
FT DOMAIN	316 394
FT DOMAIN	473 746
FT FT	
DOMAIN	403 424
TRANSNEM	425 943
FT DOMAIN	55 145
FT DOMAIN	169 303
FT DOMAIN	316 394
FT DOMAIN	473 746
FT FT	
DOMAIN	403 424
TRANSNEM	425 943
FT DOMAIN	55 145
FT DOMAIN	169 303
FT DOMAIN	316 394
FT DOMAIN	473 746
FT FT	
DOMAIN	403 424
TRANSNEM	425 943
FT DOMAIN	55 145
FT DOMAIN	169 303
FT DOMAIN	316 394
FT DOMAIN	473 746
FT FT	
DOMAIN	403 424
TRANSNEM	425 943
FT DOMAIN	55 145
FT DOMAIN	169 303
FT DOMAIN	316 394
FT DOMAIN	473 746
FT FT	
DOMAIN	403 424
TRANSNEM	425 943
FT DOMAIN	55 145
FT DOMAIN	169 303
FT DOMAIN	316 394
FT DOMAIN	473 746
FT FT	
DOMAIN	403 424
TRANSNEM	425 943
FT DOMAIN	55 145
FT DOMAIN	169 303
FT DOMAIN	316 394
FT DOMAIN	473 746
FT FT	
DOMAIN	403 424
TRANSNEM	425 943
FT DOMAIN	55 145
FT DOMAIN	169 303
FT DOMAIN	316 394
FT DOMAIN	473 746
FT FT	
DOMAIN	403 424
TRANSNEM	425 943
FT DOMAIN	55 145
FT DOMAIN	169 303
FT DOMAIN	316 394
FT DOMAIN	473 746
FT FT	
DOMAIN	

in the early formation of the chondrocytes. It seems to be required for cartilage and growth plate development.

- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.

- SUBCELLULAR LOCATION: Type I membrane protein.

- SIMILARITY: Belongs to the Tyr family of protein kinases. ROR subfamily.

- SIMILARITY: Contains 1 frizzled (Fz) domain.

- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.

- SIMILARITY: Contains 1 kringle domain.

-----

This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; ABC10384; BAA75481.1; .

HSSP; P00747; SHPG.

MGD; MGI:1347521; Ror2.

GO; GO:0001501; P:skeletal development; IMP.

InterPro; IPR000024; Fz domain.

InterPro; IPR007110; Ig-like.

InterPro; IPR003588; Ig-like.

InterPro; IPR000001; Kringle.

InterPro; IPR000719; Prot\_kinase.

InterPro; IPR001245; Tyr\_kinase.

InterPro; IPR008266; Tyr\_kinase\_AS.

Pfam; PF01392; Fz; 1.

Pfam; PF00047; Ig; 1.

Pfam; PF00051; kringle; 1.

Pfam; PF00069; kinase; 1.

PRINTS; PR00018; KRINGLE.

PRINTS; PR00109; TYRKINASE.

ProDom; PD000395; Kringle; 1.

ProDom; PD000001; Prot\_kinase; 1.

SMART; SM00408; IGC2; 1.

SMART; SM00130; KR; 1.

SMART; SM00219; TyrcK; 1.

PROSITE; PS50038; Fz; 1.

PROSITE; PS50835; Ig\_LIKE; 1.

PROSITE; PS00021; KRINGLE 1; 1.

PROSITE; PS00070; KRINGLE 2; 1.

PROSITE; PS00107; PROTEIN\_KINASE\_ATP; FALSE\_NEG.

PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.

PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.

Transferase; Tyrosine-protein kinase; ATP-binding; Receptor; Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation; Immunoglobulin domain; Developmental protein.

FT SIGNAL 1 33

FT CHAIN 34 944

FT DOMAIN 34 403

FT TRANSMEM 404 424

FT DOMAIN 425 944

FT DOMAIN 55 145

FT DOMAIN 169 303

FT DOMAIN 316 394

FT DOMAIN 473 746

FT DOMAIN 753 782

FT DOMAIN 784 857

FT DOMAIN 859 882

FT NP\_BIND 479 487

FT BINDING 507 507

FT ACT\_SITE 615 615

FT MOD\_RES 646 646

FT DISULFID 83 135

FT CARBOHYD 70 170

FT CARBOHYD 188 188

FT CARBOHYD 318 318

SEQUENCE 944 AA; 105050 MW; CD2EEBC710387A56 CRC64;

Query Match 18.7%; Score 148; DB 1; Length 944;

Best Local Similarity 36.0%; Pred. No. 5.8e-07;

Matches 40; Conservative 11; Mismatches 36; Indels 24; Gaps 8

QY 32 NC-----PKFGGQHCEIDKSKTCYEGNHGHRKASTDTWGRPCLPNNSATVLOQYH 85

DB 299 NCMRIGIPAEPLGRYH-----QCYNHGADYRGVASTTKSGHQCPW---ALQHP-H 346

QY 86 AHR---SDALQLGLGHKXNCRPNRRR-PWCYVQVGLKPLVQECNVHDC 132

DB 347 SHRLSTPEPLG-GGHAICRNFGGMEGFWCTQ-NKNRVVLCDDVPCS 395

-----

RESULT 39

PLAN\_MOUSE

ID PLAN\_MOUSE STANDARD; PRT; 812 AA.

AC P20918; O8CTS2; Q91WJ5;

DT 01-FEB-1991 (Rel. 17, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].

GN PLG.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=91184812; PubMed=2081600;

RA Degen S.J., Bell S.M., Schaefer L.A., Elliott R.W.;

RT "Characterization of the cDNA coding for mouse plasminogen and localization of the gene to mouse chromosome 17.";

RL Genomics 8:49-61(1990).

RN [2]

RP SEQUENCE FROM N.A.

RX STRAIN=129/Sv;

RA Brathwaite M.;

RT "Genomic sequence analysis in the mouse t-complex region.";

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RX TISSUE=Liver;

RC MEDLINE=22389257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh P., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Brownstein M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whitting M., Madan A., Young A.C., Shevchenko I., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickinson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.N., Krzywinski M.I., Skalek U., Smalish D.E., Scherch A., Schein J.E., Jones S.J.M., Maria M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [4]

RP SEQUENCE OF 1-16 FROM N.A.

RC STRAIN=129/SvJ; TISSUE=Liver;

RX PubMed=12149246;

RA Bannach F.G., Gutierrez A., Fowler B.J., Bugge T.H., Degen J.L., Farmer R.J., Miles L.A.;

RT "Localization of regulatory elements mediating constitutive and cytokine-stimulated plasminogen gene expression.";



DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Kremen protein 2 precursor (Kringlike-containing protein marking the eye  
 DE and the nose) (Dickkopf receptor 2).  
 GN KREMEN2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND FUNCTION.  
 RC STRAIN=C57BL/6J;  
 RX MEDLINE=22045977; PubMed=12050670;  
 RA Mao B., Wu W., Davidson G., Marhold J., Li M., Mechler B.M.,  
 RA Dellus H., Hoppe D., Stannek P., Walter C., Glinka A., Niehrs C.;  
 RT "Kremen proteins are Dickkopf receptors that regulate Wnt/beta-catenin  
 RT signalling.";  
 RL Nature 417:664-667(2002).  
 CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf  
 CC to block Wnt/beta-catenin signaling. Forms a ternary complex with  
 CC Dkk1 and LRP6 and induces rapid endocytosis and removal of the Wnt  
 CC receptor LRP6 from the plasma membrane.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).  
 CC -!- SIMILARITY: Contains 1 CUB domain.  
 CC -!- SIMILARITY: Contains 1 Kringlike domain.  
 CC -!- SIMILARITY: Contains 1 WSC domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AJ457192; CAD29805.1; -  
 DR MGD; MGI:1920266; Kremen2.  
 DR InterPro; IPR000859; CUB.  
 DR InterPro; IPR000001; Kringlike.  
 DR InterPro; IPR002889; WSC.  
 DR Pfam; PF00431; CUB; 1.  
 DR Pfam; PF00051; kringlike; 1.  
 DR Pfam; PF01822; WSC; 1.  
 DR PRINTS; PR00018; KRINGLE.  
 DR PRODOM; PD000395; Kringlike; 1.  
 DR SMART; SM00042; CUB; 1.  
 DR SMART; SM00130; KR; 1.  
 DR SMART; SM00321; WSC; 1.  
 DR PROSITE; PS01180; CUB; 1.  
 DR PROSITE; PS00021; KRINGLE 1; 1.  
 DR PROSITE; PS50070; KRINGLE 2; 1.  
 DR Wnt signaling pathway; Glycoprotein; Kringlike; Signal; Transmembrane.  
 FT SIGNAL 1 24  
 FT CHAIN 25 461  
 FT DOMAIN 25 363  
 FT TRANSMEM 364 386  
 FT DOMAIN 387 461  
 FT DOMAIN 34 118  
 FT DOMAIN 120 214  
 FT DOMAIN 218 325  
 FT CARBOHYD 48 48  
 FT CARBOHYD 221 221  
 FT CARBOHYD 243 243  
 FT CARBOHYD 350 350  
 SQ SEQUENCE 461 AA; 49170 MW; 6D58C4A3858E09DA CRC64;  
 Query Match 18.5%; Score 146.5; DB 1; Length 461;  
 Best local similarity 42.0%; Pred. No. 4e-07;  
 Matches 29; Conservative 8; Mismatches 27; Indels 5; Gaps 2;  
 QY 50 CYEKGNGHYRGK---ASTDTMGRCPLPWSNATVLQQTTHAHRSDALQLGLGKHYCNRPD 106  
 DB 35 CFQVNGADYRGHONYTGRGAGRCFLFDQIQ--QHSYSSASDPQGRWGLGAHNFCEPND 92

QY 107 NRRRPWCYV 115  
 Db 93 GDVQPCYV 101

## RESULT 41

HGFL HUMAN STANDARD; PRT; 711 AA.  
 AC P26927; Q13350; Q14870;  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-AUG-1992 (Rel. 23, Last annotation update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Hepatocyte growth factor-like protein precursor (Macrophage  
 DE stimulatory protein) (MSP) (Macrophage stimulating protein).  
 GN MST1 OR HGFL  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=92002016; PubMed=1655021;  
 RA Han S., Stuart L.A., Friesner Degen S.J.;  
 RT "Characterization of the DNFISS2 locus on human chromosome 3:  
 RT identification of a gene coding for four kringlike domains with  
 RT homology to hepatocyte growth factor.";  
 RL Biochemistry 30:9768-9780(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=93340141; PubMed=8393443;  
 RA Yoshimura T., Yuhki N., Wang M.H., Skeel A., Leonard E.J.;  
 RT "Cloning, sequencing, and expression of human macrophage stimulating  
 RT protein (MSP, MST1) confirms MSP as a member of the family of kringlike  
 RT proteins and locates the MSP gene on chromosome 3.";  
 RL J. Biol. Chem. 268:15461-15468(1993).  
 CC -!- FUNCTION: Probably has no proteolytic activity, since crucial AA  
 CC characteristic of serine proteases catalytic sites are not  
 CC conserved.  
 CC -!- PTM: MAY BE CLEAVED AFTER AA 484, TO YIELD A TWO-CHAIN MOLECULE  
 CC HELD TOGETHER BY DISULFIDE BONDS, OR TWO SEPARATE POLYPEPTIDES.  
 CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.  
 CC -!- SIMILARITY: Contains 4 kringlike domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M74178; AAA50165.1; -  
 DR EMBL; U37055; AAC50471.1; -  
 DR EMBL; L11924; AAA59872.1; -  
 DR PIR; A40331; A47136.  
 DR HSP; P00747; 2PK4.  
 DR MEROPS; S01.975; -  
 DR Genew; HGNC:7380; MST1.  
 DR MIM; 142408; -  
 DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
 DR InterPro; IPR000001; Kringlike.  
 DR InterPro; IPR003014; PAN.  
 DR InterPro; IPR003609; Pan\_app.  
 DR InterPro; IPR001254; Peptidase S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR InterPro; IPR003966; Peptidase\_S1A\_pr.  
 DR Pfam; PF00051; Kringlike; 4.  
 DR Pfam; PF00024; PAN; 1.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.



DR PRINTS; PRO0018; KRINGLE.  
DR PRINTS; PRO1505; PROTHROMBIN.  
DR ProDom; PD000395; Kringle; 4.  
DR SMART; SM00130; KR; 4.  
DR SMART; SM00473; PAN AP; 1.  
DR SMART; SM00020; TYP SPC; 1.  
DR PROSITE; PS00021; KRINGLE 1; 4.  
DR PROSITE; PS00070; KRINGLE 2; 4.  
DR PROSITE; PS00240; TRIPSPIN\_DOM; 1.  
KW Kringle; Glycoprotein; Serine protease homolog; Repeat; Signal;  
KW Polymorphism.  
FT SIGNAL 1 18  
FT CHAIN 19 711  
FT DOMAIN 32 109  
FT DOMAIN 110 186  
FT DOMAIN 191 288  
FT DOMAIN 283 361  
FT DOMAIN 370 448  
FT DOMAIN 484 711  
FT DISULFID 56 78  
FT DISULFID 60 66  
FT DISULFID 110 186  
FT DISULFID 131 189  
FT DISULFID 157 181  
FT DISULFID 191 268  
FT DISULFID 194 324  
FT DISULFID 212 251  
FT DISULFID 240 263  
FT DISULFID 283 361  
FT DISULFID 304 343  
FT DISULFID 332 355  
FT DISULFID 370 448  
FT DISULFID 391 431  
FT DISULFID 419 443  
FT DISULFID 468 588  
FT DISULFID 507 523  
FT DISULFID 602 667  
FT DISULFID 632 646  
FT DISULFID 657 685  
FT CARBOHYD 72 72  
FT CARBOHYD 296 296  
FT CARBOHYD 615 615  
FT VARIANT 13 13  
FT VARIANT 212 212  
FT VARIANT 676 676  
FT CONFLICT 623 623  
FT SEQUENCE 711 AA; 80379 MW; 596ED21F180250E4 CRC64;  
Query Match 18.2%; Score 144; DB 1; Length 711;  
Best Local Similarity 27.0%; Pred. No. 1.1e-06;  
Matches 33; Conservative 19; Mismatches 44; Indels 26; Gaps 6;  
OY 30 WC-NCPKFGGQHCIDK-----SKTCYEGKGFYRGKASTDTMGRCPLPWS 76  
Db 250 WCYTDPQIERFCDLPRGSAQPRQATVSCPRGKGEGYRGNTATTAGVPCQRWDA 309  
OY 77 ATVLQQTTHAHSDALQLGLG--KKNYCRNPDRRPMCYVQVGLKPLVQEC--MVHDC 131  
Db 310 -----QIPHQHFTEPKVACKDLRENFRCNPDGSEAPWCFT--LRPGKRAAFYQIRRC 361  
OY 132 AD 133  
Db 362 TD 363  
RESULT 42  
HGF HUMAN  
AC P14210; Q9BYL9; Q9UDU6; PRT; 728 AA.  
DT 01-JAN-1990 (Rel. 13, Created)

DT 01-AUG-1991 (Rel. 19, Last sequence update)  
DE 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hepatocyte growth factor precursor (Scatter factor) (SF)  
DE HGF OR HPTA.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91340155; PubMed=1831432;  
RA Seki T., Hagiya M., Shimonishi M., Nakamura T., Shimizu S.;  
RT "Organization of the human hepatocyte growth factor-encoding gene.";  
RN Gene 102:213-219(1991).  
[2]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Placenta;  
RX MEDLINE=89392017; PubMed=2528952;  
RA Miyazawa K., Tsubouchi H., Naka D., Takahashi K., Okigaki M.,  
RA Arakaki N., Nakayama H., Hirono S., Sakiyama O., Takahashi K.,  
RA Gonda E., Daikuhara Y., Kitamura N.;  
RT "Molecular cloning and sequence analysis of cDNA for human hepatocyte  
growth factor.";  
RN Biochem. Biophys. Res. Commun. 163:967-973(1989).  
[3]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Leukocyte;  
RX MEDLINE=91025062; PubMed=2145836;  
RA Seki T., Ihara I., Sugimura A., Shimonishi M., Mishizawa T.,  
RA Asami O., Hagiya M., Nakamura T., Shimizu S.;  
RT "Isolation and expression of cDNA for different forms of hepatocyte  
growth factor from human leukocyte.";  
RN Biochem. Biophys. Res. Commun. 172:321-327(1990).  
[4]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 55-73 AND 495-520.  
RX TISSUE=Liver;  
RX MEDLINE=90066676; PubMed=2531289;  
RA Nakamura T., Nishizawa T., Hagiya M., Seki T., Shimonishi M.,  
RA Sugimura A., Tashiro K., Shimizu S.;  
RT "Molecular cloning and expression of human hepatocyte growth factor.";  
RN Nature 342:440-443(1989).  
[5]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Embryonic fibroblast;  
RX MEDLINE=91334393; PubMed=1831266;  
RA Weidner K.M., Arakaki N., Hartmann G., Vandekerckhove J., Weingart S.,  
RA Rieder H., Fonatsch C., Tsubouchi H., Hishida T., Daikuhara Y.,  
RA Birchnieier W.;  
RT "Evidence for the identity of human scatter factor and human  
hepatocyte growth factor.";  
RN Proc. Natl. Acad. Sci. U.S.A. 88:7001-7005(1991).  
[6]  
RP SEQUENCE FROM N.A.  
RX Rieder M.J., Livingston R.J., Daniels M.R., Chung M.-W.,  
RX Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,  
RX Schackwitz W.S., Sherwood J.K., Witlak L.A., Nickerson D.A.;  
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
[7]  
RP SEQUENCE FROM N.A.  
RX Courtney L., Elliot G., Angell S.;  
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
[8]  
RP SEQUENCE OF 249-695 FROM N.A.  
RX MEDLINE=91369928; PubMed=1832556;  
RA Miyazawa K., Kitamura A., Kitamura N.;  
RT "Structural organization and the transcription initiation site of the  
human hepatocyte growth factor gene.";  
RN Biochemistry 30:9170-9176(1991).  
[9]  
RP SIGNAL SEQUENCE CLEAVAGE SITE.  
RX MEDLINE=91207365; PubMed=1826837;  
RA Yoshiyama Y., Arakaki N., Naka D., Takahashi K., Hirono S., Kondo J.,



RA Nakayama H., Gohda E., Kitamura N., Tsubouchi H., Ishii T.,  
RA Hishida T., Daikuhara Y.,  
RT "Identification of the N-terminal residue of the heavy chain of both  
RT native and recombinant human hepatocyte growth factor.";  
RL Biochem. Biophys. Res. Commun. 175:660-667(1991).  
RN [10]  
RN CARBOHYDRATE-LINKAGE SITE THR-476.  
RX MEDLINE=9312192; PubMed=1482348;  
RA Shimizu N., Hara H., Sogabe T., Sakai H., Inoue H.,  
RA Nakamura T., Shimizu S.,  
RT "Hepatocyte growth factor is linked by O-glycosylated oligosaccharide  
RT on the alpha chain.";  
RL Biochem. Biophys. Res. Commun. 189:1329-1335(1992).  
RN [11]  
RN MUTAGENESIS.  
RX MEDLINE=92331602; PubMed=1321034;  
RA Lokker N.A., Mark M.R., Luis E.A., Bennett G.L., Robbins K.A.,  
RA Baker J.B., Godowski P.J.,  
RT "Structure-function analysis of hepatocyte growth factor:  
RT identification of variants that lack mitogenic activity yet retain  
RT high affinity receptor binding.";  
RL EMBO J. 11:2503-2510(1992).  
RN [12]  
RN STRUCTURE BY NMR OF 31-127.  
RX MEDLINE=98154323; PubMed=9493272;  
RA Zhou H., Mazzulla M.J., Kaufman J.D., Stahl S.J., Wingfield P.T.,  
RA Rubin J.S., Bottaro D.P., Byrd R.A.,  
RT "The solution structure of the N-terminal domain of hepatocyte growth  
RT factor reveals a potential heparin-binding site.";  
RL Structure 6:109-116(1998).  
RN [13]  
RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 35-210.  
RX MEDLINE=99036858; PubMed=9817840;  
RA Ulsch M., Lokker N.A., Godowski P.J., de Vos A.M.,  
RT "Crystal structure of the NK1 fragment of human hepatocyte growth  
RT factor at 2.0-A resolution.";  
RL Structure 6:1383-1393(1998).  
CC -!- FUNCTION: HGF is a potent mitogen for mature parenchymal  
CC hepatocyte cells, seems to be an hepatotrophic factor, and acts  
CC as growth factor for a broad spectrum of tissues and cell types.  
CC It has no detectable protease activity.  
CC -!- SUBUNIT: Dimer of an alpha chain and a beta chain linked by a  
CC disulfide bond.  
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.  
CC -!- SIMILARITY: Contains 4 kringle domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D90334; BAA14348.1; -  
DR EMBL; D90316; BAA14348.1; JOINED.  
DR EMBL; D90319; BAA14348.1; JOINED.  
DR EMBL; D90320; BAA14348.1; JOINED.  
DR EMBL; D90322; BAA14348.1; JOINED.  
DR EMBL; D90323; BAA14348.1; JOINED.  
DR EMBL; D90324; BAA14348.1; JOINED.  
DR EMBL; D90325; BAA14348.1; JOINED.  
DR EMBL; D90326; BAA14348.1; JOINED.  
DR EMBL; D90327; BAA14348.1; JOINED.  
DR EMBL; D90328; BAA14348.1; JOINED.  
DR EMBL; D90329; BAA14348.1; JOINED.  
DR EMBL; D90330; BAA14348.1; JOINED.  
DR EMBL; D90331; BAA14348.1; JOINED.  
DR EMBL; D90332; BAA14348.1; JOINED.  
DR EMBL; D90333; BAA14348.1; JOINED.  
DR EMBL; M29145; AAA52650.1; -  
DR EMBL; M60718; AAA52648.1; -  
DR EMBL; X16323; CAA34387.1; -

DR EMBL; M73239; AAA64239.1; -  
DR EMBL; M73240; AAA64297.1; -  
DR EMBL; AY245560; AAC61091.1; -  
DR EMBL; AC004960; AAC71855.1; -  
DR EMBL; M75983; AAGS3460.1; -  
DR EMBL; M75972; AAGS3460.1; JOINED.  
DR EMBL; M75973; AAGS3460.1; JOINED.  
DR EMBL; M75974; AAGS3460.1; JOINED.  
DR EMBL; M75975; AAGS3460.1; JOINED.  
DR EMBL; M75976; AAGS3460.1; JOINED.  
DR EMBL; M75977; AAGS3460.1; JOINED.  
DR EMBL; M75978; AAGS3460.1; JOINED.  
DR EMBL; M75979; AAGS3460.1; JOINED.  
DR EMBL; M75980; AAGS3460.1; JOINED.  
DR EMBL; M75981; AAGS3460.1; JOINED.  
DR EMBL; M75982; AAGS3460.1; JOINED.  
DR PIR; JH0579; JH0579.  
DR PDB; 2HGF; 24-JUN-98.  
DR PDB; 1BHT; 18-NOV-98.  
DR PDB; 1NK1; 29-DEC-99.  
DR MEROPS; S01.976; -  
DR GlycoSuiteDB; P14210; -  
DR GenM; HGNC:4893; HGF.  
DR MIM; 142409; -  
DR GO; GO:0008083; F: growth factor activity; NAS.  
DR GO; GO:0007067; P: mitosis; NAS.  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR003014; PAN.  
DR InterPro; IPR003609; Pan\_app.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00051; kringle; 4.  
DR Pfam; PF00024; PAN; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000395; Kringle; 4.  
DR SMART; SM00130; KG; 4.  
DR SMART; SM00473; PAN\_AP; 1.  
DR SMART; SM00020; Tryp\_Spc; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 4.  
DR PROSITE; PS00070; KRINGLE\_2; 4.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR Growth factor; Kringle; Glycoprotein; Serine protease homolog; Repeat;  
DR Signal; 3D-structure; Polymorphism; Pyrrolidone carboxylic acid.  
FT SIGNAL 1 31  
FT CHAIN 32 494 HEPATOCYTE GROWTH FACTOR ALPHA CHAIN.  
FT CHAIN 495 728 HEPATOCYTE GROWTH FACTOR BETA CHAIN.  
FT MOD RES 32 32 PYRROLIDONE CARBOXYLIC ACID.  
FT DOMAIN 32 127 PAP.  
Query Match 17.8%; Score 141.5; DB 1; Length 728;  
Best Local Similarity 30.3%; Pred No. 2e-06;  
Matches 44; Conservative 12; Mismatches 64; Indels 25; Gaps 10  
Qy 2 NELHQVPSNDC--LNGGTCVSNKYFSNIHWC-----NCPKPGGO--HCEIDKSKTCYE 52  
Db 335 HEHDMTPENFKKDLRENYC-RNPDGSESPWCTTDPNIRVGYSQIPNCDMSHGQCYR 393  
Qy 53 GNGHFYRGKASTDTMGRCPLPWSNATVLOQTYHAH---RSDALQLGKHNKYNCRP-DNR 108  
Db 394 GNGKYNNGNLSQTRSGLTCSMDKN---MEDLRHIFWEPDASCL---NENYCRNPDDDA 443  
Qy 109 RRFWCYVQVGLKPLV--QECMVHDC 131  
Db 448 HGEWCYTG---NPLIPWDYCFISRC 469  
RESULT 43  
PLMN CANFA  
ID PLMN CANFA STANDARD; PRT; 333 AA.  
AC P80009;

01-NOV-1991 (Rel. 20, Created)  
01-NOV-1991 (Rel. 20, Last sequence update)  
16-OCT-2001 (Rel. 40, Last annotation update)  
DE Plasminogen (EC 3.4.21.7) (Fragment).  
PLG.  
GN Canis familiaris (Dog).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
EN NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Plasma;  
RX MEDLINE=90175323; PubMed=2626424;  
RA Schaller J., Straub C., Kaempfer U., Rickli E.E.;  
RT "Complete amino acid sequence of canine miniplasminogen.";  
RL Protein Seq. Data Anal. 2:445-450(1989).  
CC -!- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as  
CC a proteolytic factor in a variety of other processes including  
CC embryonic development, tissue remodeling, tumor invasion, and  
CC inflammation; in ovulation it weakens the walls of the Graafian  
CC follicle. It activates the urokinase-type plasminogen activator,  
CC collagenases and several complement zymogens, such as C1 and C5.  
CC It cleaves fibrin, fibronectin, thrombospondin, laminin and von  
CC Willebrand factor.  
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;  
CC higher selectivity than trypsin. Converts fibrin into soluble  
CC products.  
CC -!- ENZYME REGULATION: Converted into plasmin by plasminogen  
CC activators, both plasminogen and its activator being bound to  
CC fibrin. Activated with urokinase and high concentrations of  
CC streptokinase.  
CC -!- MISCELLANEOUS: Plasmin is inactivated by alpha-2-antiplasmin  
CC immediately after dissociation from the clot.  
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.  
CC -!- SIMILARITY: Contains at least 1 kringle domain.  
DR HSSP; P00747; 5HPC.  
DR MEROPS; S01.233; -.  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR InterPro; IPR003966; Peptidase\_S1A.  
DR Pfam; PF00051; kringle; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR PRINTS; PR01505; PROTHROMBIN.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS50070; KRINGLE\_2; 1.  
DR PROSITE; PS50240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HLS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;  
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen.  
FT CHAIN 1 103 PLASMIN HEAVY CHAIN A.  
FT CHAIN 104 333 PLASMIN LIGHT CHAIN B.  
FT DOMAIN 4 83 KRINGLE 5.  
FT DOMAIN 104 333 SERINE PROTEASE.  
FT DISULFID 4 83 BY SIMILARITY.  
FT DISULFID 25 66 BY SIMILARITY.  
FT DISULFID 54 78 BY SIMILARITY.  
FT DISULFID 90 208 INTERCHAIN (BY SIMILARITY).  
FT DISULFID 100 108 INTERCHAIN (BY SIMILARITY).  
FT DISULFID 130 146 BY SIMILARITY.  
FT DISULFID 222 289 BY SIMILARITY.  
FT DISULFID 252 268 BY SIMILARITY.  
FT DISULFID 279 307 BY SIMILARITY.  
FT ACT\_SITE 145 145 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 188 188 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT\_SITE 283 283 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT SITE 152 152 STREPTOKINASE-BINDING SITE (PROBABLE).  
FT SITE 186 186 STREPTOKINASE-BINDING SITE (PROBABLE).  
FT SITE 264 264 STREPTOKINASE-BINDING SITE (PROBABLE).  
FT SITE 277 277 SITE OF SUBSTRATE SPECIFICITY  
(BY SIMILARITY).  
SQ SEQUENCE 333 AA; 36678 MW; C8C0271B6C6AC8D4 CRC64;  
Query Match 17.8%; Score 141; DB 1; Length 333;  
Best Local Similarity 41.4%; Pred. No. 1e-06;  
Matches 36; Conservative 5; Mismatches 34; Indels 12; Gaps 5

QY 50 CYEENGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHR----SDALQLGKHNYCRNP 105  
Db 4 CMFGNGKGYRGKATVTMGIPQCEWAA-----QEPHHSIFPTPTNPQAGLEK-NYCRNP 57  
QY 106 D-NRRPWCYVQVGLKPLVQECMVHDC 131  
Db 58 DGVNGFWCYT-MNQKLFYDCVQPC 83

RESULT 44  
HGFL MOUSE STANDARD; PRT; 716 AA.  
AC P26928; 1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hepatocyte growth factor-like protein precursor (Macrophage  
DE stimulatory protein) (MSP).  
GN MST1 OR HGFL.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/c; TISSUE=Liver;  
RX MEDLINE=92002017; PubMed=1832957;  
RA Priezenr Degen S.J., Stuart L.A., Han S., Jamison C.S.;  
RT "Characterization of the mouse cdna and gene coding for a hepatocyte  
RT growth factor-like protein: expression during development.";  
RL Biochemistry 30:9781-9791(1991).  
CC -!- FUNCTION: Probably has no proteolytic activity, since crucial AA  
CC characteristic of serine proteases catalytic sites are not  
CC conserved.  
CC -!- TISSUE SPECIFICITY: Liver. Lower levels in lung, placenta and  
CC adrenal.  
CC -!- DEVELOPMENTAL STAGE: Is expressed at low levels during gestation.  
CC Just before birth the level increases dramatically and remains  
CC stable afterwards.  
CC -!- PTM: MAY BE CLEAVED AFTER AA 498, TO YIELD A TWO-CHAIN MOLECULE  
CC HELD TOGETHER BY DISULFIDE BONDS, OR CLEAVED INTO TWO SEPARATE  
CC POLYPEPTIDES.  
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.  
CC -!- SIMILARITY: Contains 4 kringle domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboratio  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
CC the European Bioinformatics Institute. There are no restrictions on it  
CC use by non-profit institutions as long as its content is in no wa  
CC modified and this statement is not removed. Usage by and for commercia  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; M74180; AAA50166.1; -.  
EMBL; M74181; AAA50167.1; -.  
PIR; A40332; A40332.  
HSSP; P00747; IRRN.  
MEROPS; S01.975; -.  
MGD; MGI:96080; Mst1.  
GO; GO:0007566; P:embryo implantation; IC.  
InterPro; IPR009003; Cys\_Ser\_trypsin.

```

DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; Pan.
DR InterPro; IPR003609; Pan app.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR003966; Peptidase_S1A_pr.
DR Pfam; PF00051; Kringle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; Kringle.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN AP; 1.
DR SMART; SM00020; TRYP SP; 1.
DR PROSITE; PS00021; Kringle_1; 4.
DR PROSITE; PS00070; Kringle_2; 4.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
KW Kringle; Glycoprotein; Serine protease homolog; Repeat; Signal.
FT SIGNAL 1 18
FT CHAIN 19 716
FT DOMAIN 19 109
FT DOMAIN 110 186
FT DOMAIN 191 268
FT DOMAIN 232 370
FT DOMAIN 379 457
FT DOMAIN 489 716
FT DISULFID 56 78
FT DISULFID 60 66
FT DISULFID 110 186
FT DISULFID 131 169
FT DISULFID 157 181
FT DISULFID 191 268
FT DISULFID 194 333
FT DISULFID 212 251
FT DISULFID 240 263
FT DISULFID 292 370
FT DISULFID 313 352
FT DISULFID 341 364
FT DISULFID 379 457
FT DISULFID 400 440
FT DISULFID 428 452
FT DISULFID 477 593
FT DISULFID 512 528
FT DISULFID 607 672
FT DISULFID 637 651
FT DISULFID 662 690
FT CARBOHYD 72 72
FT CARBOHYD 173 173
FT CARBOHYD 305 305
FT CARBOHYD 620 620
FT CONFLICT 19 19
FT SEQUENCE 716 AA; 80588 MW; BBCE025F85213ACC CRC64;

Query Match 17.5%; Score 139; DB 1; Length 716;
Best Local Similarity 39.6%; Pred. No. 3.4e-06;
Matches 38; Conservative 4; Mismatches 32; Indels 22; Gaps 7;

QY 30 WC---NCPKFGGQHCIEDKSKTCYEG-----NGHYRGKASDTDTMGRCPLWNSATVLQ 81
DB 168 WCYTHRSVRF--QSGI---KICRAVCVLCNGEDYRGVDVTESGRQCQWD-----L 217

QY 82 QYTHAH---RSDALQLGLGHKHYCRNPDRRRPCWY 114
DB 218 QHPHSHFPQPEKFLDKDL-KDNYCRNPDGSGRPWCY 252

RESULT 45
HGF_MOUSE
ID -HGF_MOUSE STANDARD; PRT; 728 AA.
AC Q08048; Q61662; Q64007;
DT 01-NOV-1995 (Rel. 32, Created)

```

```

DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor precursor (Scatter factor) (SF)
DE (Hepatopoietin A).
GN HGF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN RP SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT), AND SEQUENCE OF 496-504.
TX TISSUE=Mammary fibroblast;
RX MEDLINE=9418357; PubMed=8135822;
RA Sasaki M., Nishio M., Sasaki T., Enami J.;
RT "Identification of mouse mammary fibroblast-derived mammary growth
RL factor as hepatocyte growth factor.";
RN Biochem. Biophys. Res. Commun. 199;772-779(1994).
[2]
RN RP SEQUENCE FROM N.A.
TX TISSUE=Liver;
RX MEDLINE=94363381; PubMed=8081873;
RA Lee C.C., Kozak C.A., Yamada K.M.;
RT "Structure, genetic mapping, and expression of the mouse Hgf/scatter
RL factor gene.";
RN Cell Adhes. Commun. 1:101-111(1993).
[3]
RN RP SEQUENCE FROM N.A.
TX TISSUE=Liver;
RX MEDLINE=94060105; PubMed=8241272;
RA Liu Y., Michalopoulos G.K., Zarnegar R.;
RT "Molecular cloning and characterization of cDNA encoding mouse
RL hepatocyte growth factor.";
RN Biochim. Biophys. Acta 1216:299-303(1993).
CC -1- FUNCTION: HGF is a potent mitogen for mature parenchymal
CC hepatocyte cells, seems to be an hepatotrophic factor, and acts as
CC growth factor for a broad spectrum of tissues and cell types. It
CC has no detectable protease activity.
CC -1- SUBUNIT: Dimer of an alpha chain and a beta chain linked by a
CC disulfide bond.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=Q08048-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=Q08048-2; Sequence=VSP_005408;
CC -1- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
CC -1- SIMILARITY: Contains 4 kringle domains.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboratic
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on it
CC use by non-profit institutions as long as its content is in no wa
CC modified and this statement is not removed. Usage by and for commercia
CC entities requires a license agreement (See http://www.isb-sib.ch/announce
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; D10212; BAA01064.1; -
CC EMBL; D10213; BAA01065.1; -
CC EMBL; S71816; AAB31855.1; -
CC EMBL; X72307; CAA51054.1; ALT_INIT.
CC PIR; JC2117; A60185.
CC HSP; P14210; IBHT.
CC MEROPS; S01.982; -.
CC MGD; MGI:96079; Hgf.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR003014; PAN.
CC InterPro; IPR003609; Pan app.
CC InterPro; IPR001254; Peptidase_S1.
CC InterPro; IPR001314; Peptidase_S1A.
CC Pfam; PF00051; Kringle; 4.
CC Pfam; PF00024; PAN; 1.
CC Pfam; PF00089; trypsin; 1.

```



```
Query Match 17.5%; Score 139; DB 1; Length 728;
Best Local Similarity 29.5%; Pred. No. 3.4e-06;
Matches 44; Conservative 10; Mismatches 57; Indels 38; Gaps 11;

QY 5 HGV-PSNDC--LNGGTGVSNNKFSNIHWC-----NCPKFGGQHCEIDSK 48
DB 338 HDITENFKCDLRENYC-RNPDAESPWCFTDNIIRVGYCSQIPK-----CDVSSGQ 390
QY 49 TCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAH---RSDALQGLGKKNYCRNP 105
DB 391 DCYRGNGKYNMNLKSTRSLGTCNMWKN---MEDLHRHFWEPDASKL---TKNYCRNP 444
QY 106 -DNRPRPCYVQVGLKPLV--QECMVHDC 131
DB 445 DDAHGPWCYTG---NPLVWDYCPISRC 470

RESULT 47
PLMN PETMA STANDARD; PRT; 325 AA.
AC P33574;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Plasminogen (EC 3.4.21.7) (Fragments).
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hypercoartia;
OC Petromyzoniformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP Affolter M., Schaller J., Rickli E.E.;
RT "Isolation, characterization and partial amino acid sequence of
RT lamprey plasminogen."
RL Protein Seq. Data Anal. 5:207-211(1993).
CC -!- FUNCTION: Plasmin dissolves the fibrin of blood clots and acts as
CC a proteolytic factor in a variety of other processes including
CC embryonic development, tissue remodeling, tumor invasion,
CC and inflammation.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasminogen subfamily.
CC -!- SIMILARITY: Contains at least 2 kringle domains.
DR PIR; S33879; S33879.
DR MEROPS; P00747; SHPG.
DR MEROPS; S01.233; -.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR PRIN1S; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS02400; TRYPSIN_DOM; PARTIAL.
DR PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
DR PROSITE; PS00135; TRYPSIN_SER; PARTIAL.
KW Hydrolyase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen.
FT NON_CONS 15 16
FT NON_CONS 34 35
FT NON_CONS 44 45
FT NON_CONS 76 77
FT NON_CONS 111 112
FT NON_CONS 138 139
FT NON_CONS 158 159
FT NON_CONS 178 179
FT NON_CONS 216 217
FT NON_CONS 236 237
FT NON_CONS 268 269
FT NON_CONS 282 283

Query Match 17.2%; Score 136; DB 1; Length 325;
Best Local Similarity 30.7%; Pred. No. 3.1e-06;
Matches 31; Conservative 7; Mismatches 41; Indels 22; Gaps 3

QY 34 PKKFGGQHCEIDSKTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSALQ 93
DB 71 PQSFAG-----LTACTVGTGEGYGTAAALTVSGKACQAWASQT-----PGDVYS 115
QY 94 LGLGKKNYCRNPDRRPMYVQVGLKPLVQECMVHDCADG 134
DB 116 CQGLVSNYCRNPDEKPLWCYT-----TEYCNVPSCTGG 149

RESULT 48
THRB MOUSE STANDARD; PRT; 618 AA.
AC P19221;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Prothrombin precursor (EC 3.4.21.5).
GN F2 OR CPZ.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Liver;
RX MEDLINE=9102551; PubMed=2222810;
RA Friesner Degen S.J., Schaffer L.A., Jamison C.S., Grant S.G.,
RA Fitzgibbon J.J., Pai J.A., Chapman V.M., Elliott R.W.;
RA "Characterization of the cDNA coding for mouse prothrombin, and
RA localization of the gene on mouse chromosome 2."
RA DNA Cell Biol. 9:487-498(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Schaefer C.P., Bhat N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Max S.I., Wang J., Hsieh P.,
RA Hopkins R.F., Jordan H., Moore T., Farmer A.A., Rubin G.M., Hong L.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshlyuk S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Millaby S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallaloo D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Maman A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length
RA human and mouse cDNA sequences."
RA Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
RN [3]
RP SEQUENCE OF 384-618 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92212913; PubMed=1557383;
RA Banfield D.K., Macgillivray R.T.;
RA "Partial characterization of vertebrate prothrombin cDNAs:
RA amplification and sequence analysis of the B chain of thrombin from
RA nine different species.";
```

Proc. Natl. Acad. Sci. U.S.A. 89:2779-2783(1992).

CC -!- FUNCTION: Thrombin, which cleaves bonds after Arg and Lys,  
 CC converts fibrinogen to fibrin and activates factors V, VII, VIII,  
 CC XIII, and, in complex with thrombomodulin, protein C.  
 CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Gly; activates  
 CC fibrinogen to fibrin and releases fibrinopeptide A and B.  
 CC -!- PTM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,  
 CC RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSO-  
 CC ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES  
 CC ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY  
 CC CHARGED PHOSPHOLIPID SURFACE, WHICH IS ESSENTIAL FOR THE CONVERSION  
 CC OF PROTHROMBIN TO THROMBIN.  
 CC -!- MISCELLANEOUS: Prothrombin is activated on the surface of a  
 CC phospholipid membrane that binds the amino end of prothrombin and  
 CC factors Va and Xa in Ca-dependent interactions; factor Xa removes  
 CC the activation peptide and cleaves the remaining part into light  
 CC and heavy chains, the activation process starts slowly because  
 CC factor V itself has to be activated by the initial, small amounts  
 CC of thrombin.  
 CC -!- MISCELLANEOUS: Thrombin can itself cleave the amino terminal  
 CC fragment (fragment 1) of the prothrombin, prior to its activation  
 CC by factor Xa.  
 CC -!- SIMILARITY: Belongs to peptidase family S1.  
 CC -!- SIMILARITY: Contains 2 kringle domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL; X52308; CAA36548.1; --  
 CC DR EMBL; BC013662; AAH13662.1; --  
 CC DR EMBL; M81394; AAA40435.1; --  
 CC DR PIR; A35827; A35827.  
 CC DR HSSP; P00734; 1B7X.  
 CC DR MEROPS; S01.217; --  
 CC DR MGD; MGI:88380; F2.  
 CC DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
 CC DR InterPro; IPR002383; GLA\_blood.  
 CC DR InterPro; IPR000001; Kringle.  
 CC DR InterPro; IPR001254; Peptidase\_S1.  
 CC DR InterPro; IPR001314; Peptidase\_S1A.  
 CC DR InterPro; IPR003966; Peptidase\_S1A\_pr.  
 CC DR InterPro; IPR000294; VitK\_dep\_GLA.  
 CC DR Pfam; PF00594; gla; 1.  
 CC DR Pfam; PF00051; kringle; 2.  
 CC DR Pfam; PF00089; trypsin; 1.  
 CC DR PRINTS; PR00722; CHYMOTRYPSIN.  
 CC DR PRINTS; PR00001; GLABLOOD.  
 CC DR PRINTS; PR00018; KRINGLE.  
 CC DR PRINTS; PR01505; PROTHROMBIN.  
 CC DR ProDom; PD000395; Kringle; 2.  
 CC DR SMART; SM00069; GLA; 1.  
 CC DR SMART; SM01130; KR; 2.  
 CC DR SMART; SM00020; TRYP\_Ser; 1.  
 CC DR PROSITE; PS00011; GLU\_CARBOXYLATION; 1.  
 CC DR PROSITE; PS00021; KRINGLE\_1; 2.  
 CC DR PROSITE; PS00070; KRINGLE\_2; 2.  
 CC DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 CC DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 CC DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 CC Blood coagulation; Plasma; Calcium-binding; Glycoprotein; Repeat;  
 CC Vitamin K; Zymogen; Gamma-carboxyglutamic acid; Acute phase; Liver;  
 CC Hydrolase; Serine protease; Kringle; Signal.  
 CC SIGNAL 1 24  
 CC 25 43  
 CC 44 618  
 CC CHAIN 44 200 PROTHROMBIN.  
 CC PEPTIDE 44 200 ACTIVATION PEPTIDE (FRAGMENT 1).  
 CC PEPTIDE 201 324 ACTIVATION PEPTIDE (FRAGMENT 2).  
 CC CHAIN 325 360 THROMBIN LIGHT CHAIN (A).

FT CHAIN 361 618 THROMBIN HEAVY CHAIN (B).  
 FT DOMAIN 109 187 KRINGLE 1.  
 FT DOMAIN 215 292 KRINGLE 2.  
 FT DOMAIN 361 618 SERINE PROTEASE.  
 FT SITE 200 201 CLEAVAGE (BY THROMBIN).  
 FT SITE 324 325 CLEAVAGE (BY FACTOR XA).  
 FT SITE 360 361 CLEAVAGE (BY FACTOR XA).  
 FT ACT\_SITE 403 403 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 459 459 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 565 565 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT MOD\_RES 50 50 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 51 51 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 58 58 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 60 60 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 83 83 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 84 84 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 69 69 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 70 70 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 73 73 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT MOD\_RES 76 76 GAMMA-CARBOXYGLUTAMIC ACID.  
 FT DISULFID 61 66 BY SIMILARITY.  
 FT DISULFID 91 104 BY SIMILARITY.  
 FT DISULFID 109 187 BY SIMILARITY.  
 FT DISULFID 130 170 BY SIMILARITY.  
 FT DISULFID 158 182 BY SIMILARITY.  
 FT DISULFID 215 292 BY SIMILARITY.  
 FT DISULFID 236 276 BY SIMILARITY.  
 FT DISULFID 264 288 BY SIMILARITY.  
 FT DISULFID 333 479 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 388 404 BY SIMILARITY.  
 FT DISULFID 533 547 BY SIMILARITY.  
 FT DISULFID 561 591 BY SIMILARITY.  
 FT CARBOHYD 122 122 N-LINKED (GLCNAC. .).  
 FT CARBOHYD 144 144 N-LINKED (GLCNAC. .).  
 FT CARBOHYD 413 413 N-LINKED (GLCNAC. .).  
 FT CARBOHYD 553 553 N-LINKED (GLCNAC. .).  
 SQ SEQUENCE 618 AA; 70268 MW; B89F719A AFD601E0 CRC64;  
 Query Match 15.8%; Score 125.5; DB 1; Length 618;  
 Best Local Similarity 31.8%; Pred. No. 6.2e-05;  
 Matches 27; Conservative 13; Mismatches 40; Indels 5; Gaps 3

QY 50 CYEGNGHFYRGKASTDTMGRCPLPWN SATVLOOTYHARS DALQLGLKHNYCRNPD-NR 108  
 DB 215 CLTERGRLYQGNLAVTTLTGSPCLPWN SLPAKTL SKYQDFDPEVKL---VENFCRNPDWE 271  
 QY 109 RRPWCYVQVGLXPLVQECNVHDCAD 133  
 DB 272 EGNWCYV-AGQPDFEYCNVCEE 295

RESULT 49  
 ID THRB\_HUMAN  
 AC P00734;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Prothrombin precursor (EC 3.4.21.5) (Coagulation factor II).  
 OS Homo sapiens (Human).  
 GN Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88077877; Pubmed=2825773;  
 RA Degen S.J.F., Davie B.W.;  
 RT "Nucleotide sequence of the gene for human prothrombin.";  
 RL Biochemistry 26:6165-6177(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND VARIANT MET-165.  
 RA Rieder M.J., Armel T.Z., Carrington D.P., Chung M.-W., Lee K.L.,

RA Ozuna M., Poel C.L., Toth E.J., Yi Q., Nickerson D.A.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 8-622 FROM N.A.  
 RX MEDLINE=9233469; PubMed=6305407;  
 RA Degen S.J.F., McGillivray R.T.A., Davie E.W.;  
 RT "Characterization of the complementary deoxyribonucleic acid and gene  
 RT coding for human prothrombin.";  
 RL Biochemistry 22:2087-2097(1983).  
 RN [4]  
 RP SEQUENCE OF 44-314.  
 RX MEDLINE=77193964; PubMed=266717;  
 RA Walz D.A., Hewett-Emsmet D., Seegers W.H.;  
 RT "Amino acid sequence of human prothrombin fragments 1 and 2.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 74:1969-1972(1977).  
 RN [5]  
 RP SEQUENCE OF 315-622.  
 RX MEDLINE=77207112; PubMed=873923;  
 RA Butkowski R.J., Elin J., Downing M.R., Mann K.G.;  
 RT "Primary structure of human prothrombin 2 and alpha-thrombin.";  
 RL J. Biol. Chem. 252:4942-4957(1977).  
 RN [6]  
 RP PROCESSING.  
 RX MEDLINE=87008532; PubMed=3759958;  
 RA Rabiet M.J., Blashill A., Furie B., Furie B.C.;  
 RT "Prothrombin fragment 1 X 2 X 3, a major product of prothrombin  
 RT activation in human plasma.";  
 RL J. Biol. Chem. 261:13210-13215(1986).  
 RN [7]  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
 RX MEDLINE=90059942; PubMed=2583108;  
 RA Bode W., Mayr I., Baumann U., Huber R., Stone S.R., Hofsteenge J.;  
 RT "The refined 1.9 A crystal structure of human alpha-thrombin:  
 RT interaction with D-Phe-Pro-Arg chloromethylketone and significance of  
 RT the Tyr-Pro-Phe-Trp insertion segment.";  
 RL EMBO J. 8:3467-3475(1989).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
 RX MEDLINE=90327074; PubMed=2374926;  
 RA Rydel T.J., Ravichandran K.G., Tulinsky A., Bode W., Huber R.,  
 RT "The structure of a complex of recombinant hirudin and human alpha-  
 RT thrombin.";  
 RL Science 249:277-280(1990).  
 RN [9]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).  
 RX MEDLINE=94350942; PubMed=8071320;  
 RA Rydel T.J., Yin M., Padmanabhan K.P., Blankenship D.T., Cardin A.D.,  
 RA Correa P.E., Fenton J.W. II, Tulinsky A.;  
 RT "Crystallographic structure of human gamma-thrombin.";  
 RL J. Biol. Chem. 269:22000-22006(1994).  
 RN [10]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
 RX MEDLINE=97357286; PubMed=9214615;  
 RA van de Locht A., Bode W., Huber R., le Bonniec B.F., Stone S.R.,  
 RA Ramon C.T., Stubbs M.T.;  
 RT "The thrombin E192Q-BPTI complex reveals gross structural  
 RT rearrangements: implications for the interaction with antithrombin  
 RT and thrombomodulin.";  
 RL EMBO J. 16:2977-2984(1997).  
 RN [11]  
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 328-601.  
 RX MEDLINE=99162521; PubMed=10051558;  
 RA Guinco E.R., Caccia S., Rose T., Fuetterer K., Wakeman G., di Cera E.;  
 RT "Unexpected crucial role of residue 225 in serine proteases.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:1852-1857(1999).  
 RN [12]  
 RP VARIANT DYSPROTHROMBINEMIA CYS-314.  
 RX MEDLINE=87033739; PubMed=3771562;  
 RA Rabiet M.-J., Furie B.C., Furie B.;  
 RT "Molecular defect of prothrombin Barcelona. Substitution of cysteine  
 RT for arginine at residue 273.";  
 RL J. Biol. Chem. 261:15045-15048(1986).  
 RN [13]  
 RP VARIANT DYSPROTHROMBINEMIA ALA-509.  
 RX MEDLINE=95313001; PubMed=7792730;  
 RA Degen S.J.F., McDowell S.A., Sparks L.M., Scharer I.;  
 RT "Prothrombin Frankfurt: a dysfunctional prothrombin characterized by  
 RT substitution of Glu-466 by Ala.";  
 RL Thromb. Haemost. 73:203-209(1995).  
 RN [14]  
 RP VARIANT DYSPROTHROMBINEMIA THR-380 AND HIS-431.  
 RX MEDLINE=93043342; PubMed=1421398;  
 RA Norishita E., Saito M., Kumabaahiri I., Asakura H., Matsuda T.,  
 RA Yameguchi K.;  
 RT "Prothrombin Himi: a compound heterozygote for two dysfunctional  
 RT prothrombin molecules (Met-337-->Thr and Arg-388-->His).";  
 RL Blood 80:2275-2280(1992).  
 RN [15]  
 RP VARIANT DYSPROTHROMBINEMIA HIS-314.  
 RX MEDLINE=95169898; PubMed=7865694;  
 RA James H.L., Kim D.J., Zheng D.-Q., Girolami A.;  
 RT "Prothrombin Padua I: incomplete activation due to an amino acid  
 RT substitution at a factor Xa cleavage site.";  
 RL Blood Coagul. Fibrinolysis 5:841-844(1994).  
 RN [16]  
 RP VARIANT DYSPROTHROMBINEMIA CYS-425.  
 RX MEDLINE=89207504; PubMed=3242619;  
 RA Henriksen R.A., Mann K.G.;  
 RT "Identification of the primary structural defect in the dysfibrinogen  
 RT thrombin Quick I: substitution of cysteine for arginine-382.";  
 RL Biochemistry 27:9160-9165(1988).  
 RN [17]  
 RP VARIANT DYSPROTHROMBINEMIA VAL-601.  
 RX MEDLINE=89247398; PubMed=2719946;  
 RA Henriksen R.A., Mann K.G.;  
 RT "Substitution of valine for glycine-558 in the congenital dysfibrinogen  
 RT thrombin Quick II alters primary substrate specificity.";  
 RL Biochemistry 28:2078-2082(1989).  
 RN [18]  
 RP VARIANT DYSPROTHROMBINEMIA ALA-509.  
 RX MEDLINE=92378975; PubMed=1354985;  
 RA Miyata T., Aruga R., Umeyama H., Bezeaud A., Guillin M.-C.,  
 RA Iwanaga S.;  
 RT "Prothrombin Salakta: substitution of glutamic acid-466 by alanine  
 RT reduces the fibrinogen clotting activity and the esterase activity.";  
 RL Biochemistry 31:7457-7462(1992).  
 RN [19]  
 RP VARIANT DYSPROTHROMBINEMIA TRP-461.  
 RX MEDLINE=87185407; PubMed=3567158;  
 RA Miyata T., Morita T., Inomoto T., Kawauchi S., Shirakami A.,  
 RA Iwanaga S.;  
 RT "Prothrombin Tokushima, a replacement of arginine-418 by tryptophan  
 RT that impairs the fibrinogen clotting activity of derived thrombin  
 RT Tokushima.";  
 RL Biochemistry 26:1117-1122(1987).  
 RN [20]  
 RP VARIANT DYSPROTHROMBINEMIA TRP-461.  
 RX MEDLINE=87101511; PubMed=3801671;  
 RA Inomoto T., Shirakami A., Kawauchi S., Shigeakiyo T., Saito S.,  
 RA Miyoshi K., Morita T., Iwanaga S.;  
 RT "Prothrombin Tokushima: characterization of dysfunctional thrombin  
 RT derived from a variant of human prothrombin.";  
 RL Blood 69:565-569(1987).  
 RN [21]  
 RP VARIANT DYSPROTHROMBINEMIA TRP-461.  
 RX MEDLINE=92256895; PubMed=1349838;  
 RA Iwahana H., Yoshimoto K., Shigeakiyo T., Shirakami A., Saito S.,  
 RA Itakura M.;  
 RT "Detection of a single base substitution of the gene for prothrombin  
 RT Tokushima. The application of PCR-SSCP for the genetic and molecular  
 RT analysis of dysprothrombinemia.";  
 RL Int. J. Hematol. 55:93-100(1992).  
 RN [22]  
 RP VARIANT DYSPROTHROMBINEMIA LYS-200.  
 RX MEDLINE=83204687; PubMed=6405779;



RA Board P.G., Shaw D.C.;  
RT "Determination of the amino acid substitution in human prothrombin  
RT type 3 (157 Glu leads to Lys) and the localization of a third  
RT thrombin cleavage site";  
RL Br. J. Haematol. 54:245-254(1983).  
RN [23]  
RX VARIANTS MET-165 AND THR-386.  
RY MEDLINE=99318093; PubMed=10391209;  
RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemes J., Ziaugra L.,  
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,  
RA Lander E.S.;  
RT "Characterization of single-nucleotide polymorphisms in coding regions  
RT of human genes";  
RL Nat. Genet. 22:231-238(1999).  
RN [24]  
RP ERRATUM.  
RA Cargill M., Altshuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemes J., Ziaugra L.,  
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,  
RA Lander E.S.;  
RL Nat. Genet. 23:373-373(1999).  
CC -!- FUNCTION: Thrombin, which cleaves bonds after Arg and Lys,  
CC converts fibrinogen to fibrin and activates factors V, VII, VIII,  
CC XIII, and in complex with thrombomodulin, protein C.  
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Gly; activates  
CC fibrinogen to fibrin and releases fibrinopeptide A and B.  
CC -!- SUBCELLULAR LOCATION: Extracellular.  
CC -!- TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.  
CC -!- PTM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS,  
CC RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOMAL  
CC ENZYME. THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES  
CC ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY  
CC CHARGED PHOSPHOLIPID SURFACE, WHICH IS ESSENTIAL FOR THE CONVERSION

Query Match 15.6%; Score 124; DB 1; Length 622;  
Best Local Similarity 38.8%; Pred. No. 8.7e-05;  
Matches 26; Conservative 8; Mismatches 29; Indels 4; Gaps 2;  
QY 50 CYESGNGHFRGKASTDTWGRPCLPNSATVLTQTHAHSALQGLGKHYCNRPD-NR 108  
DB 213 CVPDRGQYQGRVAVTTHGLPLCLAWASAKALSKHDFNSAVQL---VENFCNRPDGE 269  
QY 109 RRPWCYV 115  
DB 270 EGVWCYV 276  
[1]  
RESULT 50  
ROR2 DROME  
ID ROR2 DROME STANDARD; PRT; 724 AA.  
AC Q9V6K3; O02001; O96391; Q9TVH9;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Tyrosine-protein kinase transmembrane receptor Ror2 precursor  
DE (EC 2.7.1.112) (Neurospecific receptor tyrosine kinase).  
GN NRK OR ROR2 OR CG4007.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RN SEQUENCE FROM N.A.  
RA Frick K.J., Scott M.J.;  
RL Submitted (DEC-1997) to the EMBL/GenBank/DBSJ databases.  
RN [2]  
RN SEQUENCE FROM N.A.  
RP STRAIN=Berkeley;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celnikner S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,

George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.B.G.,  
Abril J.F., Agbayani A., An H.-J., Andrews-Fannkoch C., Baldwin D.,  
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
Beeson K.Y., Beron P.V., Berman B.P., Bhandari D., Bolshakov S.,  
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
Dodonson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
Jalali M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,  
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
Markulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
Nelson D.R., Nelson K.A., Nixon K., Nusekern D.R., Pacleb J.M.,  
Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
Wang Z.-Y., Wassman D.A., Weinstock G.M., Weissenbach J.,  
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster";  
RL Science 287:2185-2195(2000).  
RN [3]  
RP REVISIONS.  
RX MEDLINE=22426069; PubMed=12537572;  
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
RA Battencourt B.R., Celnikner S.E., de Grey A.D.N.J., Drysdale R.A.,  
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
RA Lewis S.E.;  
RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
RT systematic review";  
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
RN [4]  
RP SEQUENCE OF 17-724 FROM N.A., AND CHARACTERIZATION.  
RC TISSUE=Imaginal disks.  
RX MEDLINE=97277331; PubMed=9115253;  
RA Oishi I., Sugiyama S., Liu Z.J., Yamamura H., Nishida Y., Minami Y.;  
RT "A novel Drosophila receptor tyrosine kinase expressed specifically  
RT in the nervous system. Unique structural features and implication in  
RT developmental signaling";  
RN J. Biol. Chem. 272:11916-11923(1997).  
RN [5]  
RP SEQUENCE OF 586-638 FROM N.A.  
RX MEDLINE=98401146; PubMed=9731193;  
RA Oates A.C., Wollberg P., Achen M.G., Wilks A.F.;  
RT "Sampling the genomic pool of protein tyrosine kinase genes using the  
RT polymerase chain reaction with genomic DNA";  
RL Biochem. Biophys. Res. Commun. 249:660-667(1998).  
CC -!- FUNCTION: Tyrosine-protein kinase receptor that functions during  
CC early stages of neuronal development (By similarity).  
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).  
CC -!- TISSUE SPECIFICITY: Expressed in neural cell lineage from  
CC embryonic stage 11 onwards, resulting in expression in the brain  
CC and ventral nerve cord at the end of embryogenesis.



CC -!- DEVELOPMENTAL STAGE: Expressed at high levels in embryos and  
CC larvae, low levels in adults and pupae show maximal expression.  
CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. ROR  
CC subfamily.  
CC -!- SIMILARITY: Contains 1 frizzled (FZ) domain.  
CC -!- SIMILARITY: Contains 1 kringle domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF037164; AAD02091.1; -  
CC EMBL; AE003819; AAF58420.2; -  
CC EMBL; AB001420; BAA20134.1; -  
CC EMBL; AJ002920; CAA05755.1; -  
CC HSSP; P11362; 1FGK.  
CC FlyBase; Fgn0020391; Nrk.  
CC GO; GO:0004713; P:protein-tyrosine kinase activity; IDA.  
CC GO; GO:0006468; P:protein amino acid phosphorylation; IDA.  
CC GO; GO:0007165; P:signal transduction; IDA.  
CC InterPro; IPR000024; Fz\_domain.  
CC InterPro; IPR000001; Kringle.  
CC InterPro; IPR000719; Prot\_kinase.  
CC InterPro; IPR001245; Tyr\_Dkinase.  
CC InterPro; IPR008266; Tyr\_Dkinase\_AS.  
CC Pfam; PF01392; Fz; 1.  
CC Pfam; PF00051; kringle; 1.  
CC Pfam; PF00069; pkinase; 1.  
CC PRINTS; PR00018; KRINGLE.  
CC PRINTS; PR00109; TYRKINASE.  
CC ProDom; PD000395; Kringle; 1.  
CC ProDom; PD000001; Prot\_kinase; 1.  
CC PROSITE; PS00038; FZ; 1.  
CC PROSITE; PS00021; KRINGLE\_1; 1.  
CC PROSITE; PS00070; KRINGLE\_2; 1.  
CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
CC PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
CC PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
CC Transferrase; Kinase; Tyrosine-protein kinase; ATP-binding; Receptor;  
KW Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation;  
KW Developmental protein.  
FT SIGNAL 1 41  
FT CHAIN 42 724  
FT  
FT  
FT DOMAIN 42 322  
FT TRANSMEM 323 343  
FT DOMAIN 344 724  
FT DOMAIN 441 711  
FT NP\_BIND 447 455  
FT BINDING 475 475  
FT ACT\_SITE 580 580  
FT MOD\_RES 606 606  
FT MOD\_RES 610 610  
FT MOD\_RES 611 611  
FT CONFLICT 51 51  
FT CONFLICT 192 192  
FT CONFLICT 306 306  
FT CONFLICT 387 387  
FT CONFLICT 391 391  
FT CONFLICT 547 547  
FT CONFLICT 706 724  
SQ SEQUENCE 724 AA; 81836 MW; 0FFB95E1F7B4F6A26 CRC64;

Query Match 15.2%; Score 120.5; DB 1; Length 724;  
Best Local Similarity 35.8%; Pred. No. 0.00022;  
Matches 34; Conservative 5; Mismatches 45; Indels 11; Gaps 5;

Qy 43 EIDKSKTCYEGNGHFRGKASTDTMGRPCLPWNSATVLQOTYHAHSDAL---QLGLGH 99

Db 212 ESEVSYDCRNGRFTYGTMTVSKSGIPQORWDT-----QYPKHQPPPLVTHQL-LEGE 265  
Qy 100 NYCENFDNR-REPWCYVQVGLKPLVQECMVHDCAD 133  
Db 266 NYCENAGGEBPHWCYT-VDESVMWQHCDIPMCPD 299

Search completed: May 25, 2004, 14:53:49  
Job time : 6.20596 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2004, 14:44:05 ; Search time 22.362 Seconds  
(without alignments)  
1904.795 Million cell updates/sec

Title: US-09-880-503-4

Perfect score: 793

Sequence: 1 SNELHQVSNCDLNGTGV.....QVGLKPLVQECWVHDCADGK 135

Scoring table: BLOSOM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database :

SPTREMBL 25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriaph:\*
- 17: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	778	98.1	154	4 Q96SE8	Q96SE8 homo sapien
2	602	75.9	433	6 Q8MIL0	Q8MIL0 oryctolagus
3	602	75.9	433	6 Q8MHY7	Q8MHY7 oryctolagus
4	550	69.4	157	6 Q9TVA8	Q9TVA8 bos taurus
5	548	69.1	221	11 Q8G6L2	Q8G6L2 mus musculus
6	328.5	41.4	291	4 Q7Z7N2	Q7Z7N2 homo sapien
7	328.5	41.4	562	4 Q86YK8	Q86YK8 homo sapien
8	322.5	40.7	516	4 Q9BU99	Q9BU99 homo sapien
9	313	39.5	214	6 Q9XT70	Q9XT70 oryctolagus
10	310.5	39.2	562	6 Q8SQ23	Q8SQ23 sus scrofa
11	301.5	38.0	564	6 Q8MKB1	Q8MKB1 oryctolagus
12	267.5	33.4	517	11 Q8KOD2	Q8KOD2 mus musculus
13	257.5	32.5	560	4 Q14520	Q14520 homo sapien
14	257	32.4	616	6 Q97507	Q97507 sus scrofa
15	252.5	31.8	540	13 Q800Y7	Q800Y7 meleagris g
16	250.5	31.6	653	11 Q8VCS4	Q8VCS4 mus musculus

17	249	31.4	597	11	Q35727	Q35727 mus musculus
18	249	31.4	609	11	Q80YC5	Q80YC5 mus musculus
19	233	29.4	615	4	Q81Z25	Q81Z25 homo sapien
20	229.5	28.9	395	4	Q9B2W1	Q9B2W1 homo sapien
21	218.5	27.6	80	4	Q8NG20	Q8NG20 homo sapien
22	197.5	24.9	202	13	Q90675	Q90675 gallus gall
23	174	21.9	421	13	Q8AXX3	Q8AXX3 xenopus lae
24	170.5	21.5	947	13	Q8AXY6	Q8AXY6 gallus gall
25	166	20.9	716	13	Q91691	Q91691 xenopus lae
26	165.5	20.9	359	6	Q8WNR1	Q8WNR1 canis fami
27	165	20.8	812	11	Q9RCW3	Q9RCW3 rattus norv
28	158.5	20.0	393	4	Q9B8B6	Q9B8B6 homo sapien
29	158.5	20.0	937	11	Q8BNP9	Q8BNP9 mus musculus
30	158.5	20.0	937	11	Q8BG10	Q8BG10 mus musculus
31	158	19.9	454	6	Q46506	Q46506 papio hamad
32	156	19.7	113	4	Q9UIR5	Q9UIR5 homo sapien
33	156	19.7	759	11	Q7TP84	Q7TP84 rattus norv
34	155.5	19.6	263	13	Q7SXB3	Q7SXB3 brachydanic
35	154	19.4	806	6	Q18783	Q18783 macropus eu
36	154	19.4	810	4	Q15146	Q15146 homo sapien
37	153	19.3	113	4	Q9UIR7	Q9UIR7 homo sapien
38	152.5	19.2	717	13	P70006	P70006 xenopus lae
39	151.5	19.1	648	4	Q9H1V4	Q9H1V4 homo sapien
40	150.5	19.0	567	4	Q13208	Q13208 homo sapien
41	150.5	19.0	930	13	Q8AV69	Q8AV69 xenopus lae
42	149	18.8	709	13	Q7ZTN9	Q7ZTN9 xenopus lae
43	149	18.8	801	11	Q8KQ8	Q8KQ8 mus musculus
44	149	18.8	944	11	Q8C3W2	Q8C3W2 mus musculus
45	149	18.8	944	11	Q8BSP6	Q8BSP6 mus musculus
46	148	18.7	105	4	Q9UIR8	Q9UIR8 homo sapien
47	147	18.5	113	4	Q9UIR6	Q9UIR6 homo sapien
48	146.5	18.5	704	13	Q90865	Q90865 gallus gall
49	146.5	18.5	709	13	Q902N6	Q902N6 brachydanic
50	144.5	18.2	902	5	Q9BLY1	Q9BLY1 caenorhabdi
51	144.5	18.2	928	5	Q17576	Q17576 caenorhabdi
52	142	17.9	948	5	Q9YIY6	Q9YIY6 ephydaria f
53	141	17.8	215	13	Q42341	Q42341 gallus gall
54	141	17.8	710	13	Q91402	Q91402 xenopus. he
55	141	17.8	726	13	Q90978	Q90978 gallus gall
56	140	17.7	132	4	Q16609	Q16609 homo sapien
57	139	17.5	716	11	Q91XG8	Q91XG8 mus musculus
58	139	17.5	728	6	Q9BH09	Q9BH09 felis silve
59	139	17.5	728	11	Q8C9G5	Q8C9G5 mus musculus
60	139	17.5	730	6	Q867B7	Q867B7 canis fami
61	138	17.4	716	11	P70521	P70521 rattus norv
62	138	17.4	1145	5	Q9BKL8	Q9BKL8 aplysia cal
63	137.5	17.3	313	13	Q9PU78	Q9PU78 crocodylus
64	136.5	17.2	2869	6	Q28398	Q28398 erinaceus e
65	135.5	17.1	553	13	Q7Z292	Q7Z292 brachydanic
66	134	16.9	211	11	Q55027	Q55027 mus musculus
67	132.5	16.7	213	11	Q811Z3	Q811Z3 mus musculus
68	132.5	16.7	264	11	Q811Z2	Q811Z2 mus musculus
69	132.5	16.7	264	11	Q7TMJ8	Q7TMJ8 mus musculus
70	131	16.5	95	4	Q8N696	Q8N696 homo sapien
71	131	16.5	145	6	Q28911	Q28911 macaca fasc
72	131	16.5	234	4	Q86YW2	Q86YW2 homo sapien
73	131	16.5	263	4	Q00318	Q00318 homo sapien
74	131	16.5	263	4	Q8NCJ9	Q8NCJ9 homo sapien
75	131	16.5	263	4	Q96FE7	Q96FE7 homo sapien

#### ALIGNMENTS

RESULT 1

Q96SE8

ID Q96SE8 PRELIMINARY; PRT; 154 AA.

AC Q96SE8  
DT 01-DEC-2001 (TRENBLrel. 19, Created)  
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
DE Urokinase-type plasminogen activator amino-terminal fragment.  
GN ATF.



```

DR EMBL; AY029517; AAK40239.1; --
DR EMBL; AB087224; BAC02685.1; --
DR GO; GO:0004283; F:chymotrypsin activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR003003; Cys Ser trypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR008293; Pept_S1A_UPA.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR PIRSF; PIRSF001144; Urk plas act; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48444 MW; 6DD35A371010A6EE CRC64;

Query Match 75.9%; Score 602; DB 6; Length 433;
Best Local Similarity 76.6%; Pred. No. 2.5e-57;
Matches 105; Conservative 10; Mismatches 20; Indels 2; Gaps 1;

QY 1 SNEHQV--PNCCLNGGTCVSNKYFSNIHWCNPKKFGQHCIDKSKTCYEGNGHFY 58
DB 21 SHELGVSDASNCCLNGGTCVTKYFSNIWRCNPKKFGQHCIDTLCYHGDSY 80

QY 59 RKASTDTHGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQGLKPLVQECMVHDC 118
DB 81 RGAQNTDMRPCLAWNSANVLTXYHAHRSALQGLGKHNYCRNPNRRPWCYVQGLKPLVQECMVHDC 140

QY 119 LKPLVQECMVHDCADGK 135
DB 141 LKQLIQECKVHDCSSGK 157

```

```

RESULT 4
Q9TVAS8
ID Q9TVAS8 PRELIMINARY; PRT; 157 AA.
AC Q9TVAS8
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DE Urokinase plasminogen activator (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RA MEDLINE=21071388; PubMed=11204721;
RT "Coordinate expression of matrix-degrading proteinases and their
activators and inhibitors in bovine skeletal muscle."
RL J. Anim. Sci. 79:94-107(2001).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF144761; AAD30301.1; --
DR HSP; P00749; IURK.
DR GO; GO:0016301; F:kinase activity; IEA.
DR InterPro; IPR006209; EGF_like.

```

```

DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kinase; Kringle.
FT NON_TER 1 157
FT NON_TER 157 157
SQ SEQUENCE 157 AA; 17858 MW; A768D6C72C1FBFB7 CRC64;

Query Match 69.4%; Score 550; DB 6; Length 157;
Best Local Similarity 75.6%; Pred. No. 3.9e-52;
Matches 93; Conservative 11; Mismatches 19; Indels 0; Gaps 0

QY 13 CLNGGTCVSNKYFSNIHWCNPKKFGQHCIDKSKTCYEGNGHFYRKASTDTNRPCL 72
DB 1 CLNGGKCVTKYFSNIQRCSPKPKFGQHCIDTSTKYCQNGHSYRKANRDLSGRPCL 60

QY 73 PWSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQGLKPLVQECMVHDC 132
DB 61 AWDSPTVLLKMYHAHRSDAIQGLGKHNYCRNPNRRPWCYVQGLKPLVQECMVHDC 120

QY 133 DGK 135
DB 121 VGK 123

RESULT 5
Q8C6L2
ID Q8C6L2 PRELIMINARY; PRT; 231 AA.
AC Q8C6L2;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Plasminogen activator (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK054349; BAC35743.1; --
DR PIR; PT0534; PT0534.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
FT NON_TER 231 231
SQ SEQUENCE 231 AA; 25510 MW; 25E8980A682737F2 CRC64;

Query Match 69.1%; Score 548; DB 11; Length 231;
Best Local Similarity 71.7%; Pred. No. 9.9e-52;
Matches 91; Conservative 13; Mismatches 23; Indels 0; Gaps 0

```

QY 9 SNCLNGGTGVSNKYFSNIHWCNPKFGQHQHCEIDSKTCYEGNGHFGYRGKASTDTMG 68  
Db 30 SNGCQNGGVSVSYKIFRIRRCSPRFQGEHCEIDASKTCYHNGDSYRGKANTDTKG 69  
QY 69 RPLCPWNSATVLOQTYHAHRSALQGLGKHNCRNPNDRRPPWCYVQVGLKPLVQECWV 128  
Db 90 RPLAWNAPVLOQTYHAHRSALQGLGKHNCRNPNDRRPPWCYVQVGLKPLVQECWV 149  
QY 129 HDCADGK 135  
Db 150 HDCSLSK 156  
RESULT 6  
ID Q7Z7N2 PRELIMINARY; PRT; 291 AA.  
AC Q7Z7N2;  
DT 01-OCT-2003 (Tremblrel. 25, Created)  
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Plasmidogen activator, tissue type isoform 2.  
GN PLAT.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Rieder M.J., Carrington D.P., da Ponte S.H., Hastings N.C.,  
RA Ahearn M.O., Kuldane S.A., Rajkumar N., Toth E.J., Yi Q.,  
RA Nickerson D.A.; to the EMBL/GenBank/DBJ databases.  
RL Submitted (May-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY291060; AAF34246.1; -  
SQ SEQUENCE 291 AA; 32191 MW; 874E38C52F50EF1D CRC64;  
Query Match 41.4%; Score 328.5; DB 4; Length 291;  
Best Local Similarity 46.3%; Pred No. 1.1e-27;  
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;  
QY 3 ELHQVP-SNCD---CLNGTGVSNKYFSNIHWCNPKFGQHQHCEIDSKTCYEGNGHFGY 58  
Db 77 QCHSVFVKSCSPRCFNGTCCQALYFSD-F-VQCPEGFAGKCEIDTRATCYEDQGISY 135  
QY 59 RKGASTDTMGRLCPWNSATVLOQTYHAHRSALQGLGKHNCRNPNDRRPPWCYVQV 118  
Db 136 RGTWSTAGSAGCTWNSALQAQKPYSGRRPDAIRLGLGNHNYCRNPRDRSKPCYVFK 195  
QY 119 LKPLVQECMVHDCADG 134  
Db 196 GKYSEFCSTPACSEG 211  
RESULT 7  
ID Q86YK8 PRELIMINARY; PRT; 562 AA.  
AC Q86YK8;  
DT 01-JUN-2003 (Tremblrel. 24, Created)  
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Tissue plasmidogen activator.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Liu Y., Xu L., Zeng Y., He X.;  
RT "CDNA of tissue plasmidogen activator."  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY221101; AAO34406.1; -  
DR GO; GO:0005576; C:extracellular; IEA.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.

DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:000508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR009003; Cys Ser trypsin.  
DR InterPro; IPR006209; EGF like.  
DR InterPro; IPR000083; Fibrinctn1.  
DR InterPro; IPR000001; Kringl.  
DR InterPro; IPR001254; peptidase S1.  
DR InterPro; IPR001314; peptidase\_S1A.  
DR Pfam; PF00008; EGF; 1.  
DR Pfam; PF00039; fnl; 1.  
DR Pfam; PF00051; kringl; 2.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000395; Kringl; 2.  
DR SMART; SM00058; FNI; 1.  
DR SMART; SM00130; KR; 2.  
DR SMART; SM00020; TYP\_SPC; 1.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS01186; EGF\_2; 1.  
DR PROSITE; PS01253; FIBRONECTIN\_1; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 2.  
DR PROSITE; PS00070; KRINGLE\_2; 2.  
DR PROSITE; PS02440; TRYPsin\_DOM; 1.  
DR PROSITE; PS00134; TRYPsin\_HIS; 1.  
DR PROSITE; PS00135; TRYPsin\_SER; 1.  
SQ SEQUENCE 562 AA; 62902 MW; 837D98392F6EDD1F CRC64;  
Query Match 41.4%; Score 328.5; DB 4; Length 562;  
Best Local Similarity 46.3%; Pred No. 2.3e-27;  
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;  
QY 3 ELHQVP-SNCD---CLNGTGVSNKYFSNIHWCNPKFGQHQHCEIDSKTCYEGNGHFGY 58  
Db 77 QCHSVFVKSCSPRCFNGTCCQALYFSD-F-VQCPEGFAGKCEIDTRATCYEDQGISY 135  
QY 59 RKGASTDTMGRLCPWNSATVLOQTYHAHRSALQGLGKHNCRNPNDRRPPWCYVQV 118  
Db 136 RGTWSTAGSAGCTWNSALQAQKPYSGRRPDAIRLGLGNHNYCRNPRDRSKPCYVFK 195  
QY 119 LKPLVQECMVHDCADG 134  
Db 196 GKYSEFCSTPACSEG 211  
RESULT 8  
ID Q9BU99 PRELIMINARY; PRT; 516 AA.  
AC Q9BU99;  
DT 01-JUN-2001 (Tremblrel. 17, Created)  
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Similar to plasmidogen activator, tissue.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC Tissue-Skin;  
RA Strauberg R.;  
RL Submitted (Feb-2001) to the EMBL/GenBank/DBJ databases.  
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -I- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.  
DR EMBL; BC002795; AA02795.1; -  
DR HSP; P00750; 1A5H.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0005233; F:peptidase activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR009003; Cys Ser trypsin.  
DR InterPro; IPR006209; EGF like.  
DR InterPro; IPR006210; IEGF.

```
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00051; Kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00186; EGF_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 516 AA; 57370 MW; BAB31901FDC96800 CRC64;

Query Match 40.7%; Score 322.5; DB 4; Length 516;
Best Local Similarity 44.4%; Pred. No. 9, 5e-27;
Matches 64; Conservative 13; Mismatches 56; Indels 11; Gaps 2;

QY 1 SNEH-----QVPSNCDLNGTCVKNYFSNIHWCNCPKFGGQHCEIDKSKTC 50
DB 23 SGEIHFARFGRARSYQGCSEPCFNGTCQALYFDF-VQCPEGFAGKCCEDTRATC 81
QY 51 YEGNGHFYRGKASTDTMGRPLCPMNSATVLOQTYHAHSDALQGLGKHNYCRNPDNR 110
DB 82 YEDQISVIGTWSTAESGACINWSSALAQKPYSGRRPDALRLGLGHNHCRNPDNR 141
QY 111 PWCYVQVGLKPLVQRCVHDCADG 134
DB 142 PWCYVFKAGYSSEFCSTPACSEG 165

RESULT 9
Q9XT70 PRELIMINARY; PRT; 214 AA.
AC Q9XT70;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Urokinase-type plasminogen activator (Fragment).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Yan J., Idell S.;
RT "Partial mRNA of rabbit uPA.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF097647; AAD39351.1; -.
DR HSSP; P00749; 1EUN.
DR MEROPS; S01.231; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00051; kringle; 1.

Q8SQ23 PRELIMINARY; PRT; 562 AA.
AC Q8SQ23;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE T-plasminogen activator.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Enamel organ;
RA Ding Y., Xue J., Bartlett J.D.;
RT "T-plasminogen activator in tooth tissues.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AF364605; AAM00297.1; -.
DR HSSP; P00761; 1AN1.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR006203; EGF-like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 2.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
```

```

DR PROSITE: PS00021; KRINGLE 1; 2.
DR PROSITE: PS50070; KRINGLE 2; 2.
DR PROSITE: PS50240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
DR EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 562 AA; 63468 MW; F95E5B4C77CB101B8 CRG64;

Query Match 39.2%; Score 310.5; DB 6; Length 562;
Best Local Similarity 43.8%; Pred. No. 2.1e-25;
Matches 60; Conservative 14; Mismatches 58; Indels 5; Gaps 3;

QY 3 ELHQP-V-SNCD---CLNGGTCVSNKYFSTNTWNCNCPKKGQGHCEIDKSKTCYEGNGHFY 58
Db 77 QCHSVFVKSCEPFCFGTGLQAIYFSDF-VQCPVGVIGRCEIDARATCYEQGITY 135
QY 59 RKGASTDTWGRPCLPWNSATVLOQTYHAHSDALQLGLGHNHYCRNPDNRRFPWCYVQG 118
Db 136 RGTWSTTESGAECVNWNTSGLASPNYGRNPDPVKVLGLGHNHYCRNPDKDSKPCYIIFKA 195
QY 119 LKPLVQECWVHDCADGK 135
Db 196 EKYSPDFCGSTPACTKEK 212
QY Db

```

```

RESULT 11
Q8MBK1 Q8MBK1 PRELIMINARY; PRT; 564 AA.
ID Q8MBK1
AC Q8MBK1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE tissue-type plasminogen activator.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
CC NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RF Sugiki M., Yoshida E., Anai K., Maruyama M.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AY029518; AAK40240.1;
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001234; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 2.
DR SMART; SMC00181; EGF; 1.
DR SMART; SMC00058; FN1; 1.
DR SMART; SMC00130; KR; 2.
DR SMART; SMC0020; TRYD_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS50070; KRINGLE_2; 2.

```

DR PROSITE; PSS0240; TRYPSIN\_DOM; 1.  
DR PROSITE; PSS0134; TRYPSIN\_HIS; 1.  
DR PROSITE; PSS0135; TRYPSIN\_SER; 1.  
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;  
KW Serine protease.  
SQ SEQUENCE 564 AA; 62726 MW; 459D8BAC6D4A937C CRC64;

Query Match 38.0%; Score 301.5; DB 6; Length 564;  
Best Local Similarity 48.7%; Pred.No.2.1e-24;  
Matches 56; Conservative 13; Mismatches 41; Indels 5; Gaps 3

QY 5 HQVP-SNCD---CLNGGTCVSNKYFSNIHWCNPKFGGQHCIDKSKTKCYGNGHFYRG 60  
Db 80 HSPVQSCSEPRCLNGGTCQALYFSDP-VQCPEGFVGKRCVEDTRACYEDRGIGYRG 138  
QY 61 KASDTMGRCPCLPNWSATVLQOTVHAHRSALQLGKGNKNCNPNRRPWCYV 115  
Db 139 TWSVTESGAQCWNWSSWLAKPKYSGRKPALRLGLGNHNYCNPRDPKPCYV 193

RESULT 12

Q8K0D2 PRELIMINARY; PRT; 517 AA.

AC Q8K0D2;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_Taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RA Strausberg R.;  
EL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO PERTIDASE FAMILY S1.  
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.

DR ENBL; BC031775; AAH31775.1; -.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR000742; EGF\_2.  
DR InterPro; IPR001881; EGF Ca.  
DR InterPro; IPR006209; EGF like.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR001254; Peptidase S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00008; EGF; 2.  
DR Pfam; PF00051; kringle; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00181; EGF; 3.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00020; Tryp\_Spc; 1.  
DR PROSITE; PS00022; EGF 1; 3.  
DR PROSITE; PS01186; EGF 2; 2.  
DR PROSITE; PS00021; KRINGLE 1; 1.  
DR PROSITE; PS00070; KRINGLE 2; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;  
KW Kringle; Protease; Serine protease.  
SQ SEQUENCE 517 AA; 57326 MW; 3855A4203A5EA59 CRC64;





```
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 616 AA; 68012 MW; 4CSFE3D71EBBDA9 CRC64;

Query Match 32.4%; Score 257; DB 6; Length 616;
Best Local Similarity 40.9%; Pred. No. 1.6e-19;
Matches 52; Conservative 14; Mismatches 53; Indels 8; Gaps 3;

QY 6 QVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRGKASTD 65
Db 176 QVGVSTNPCLNGSGCLQTE---GHLRCRCPTAGRLCDVDLXERCYSRGLSYRGMAOTT 232
QY 66 TMRGRLPWNASATVLOQTY-HAHRSDALQLGLGKHNYCRNPNRRRPPRCYQVGLKPLVQ 124
Db 233 LSGAPQCPWAS-----EATYWNMTAEQALNWLGDHAFCEPNPDTRPWCFFVRGDLQSNQ 288
QY 125 ECMWHDC 131
Db 289 YCRLARC 295

RESULT 15
Q800Y7 PRELIMINARY; PRT; 540 AA.
AC Q800Y7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hepatocyte growth factor activator (Fragment).
OS Meleagris gallopavo (Common turkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Meleagris.
OX NCBI_TaxID=9103;
RN [1]
RP MEDLINE=2212796; PubMed=12128063;
RA Holsberger D.R., Becker A.E., Thurston R.J., Rice C.D.;
RT "Expression of a hepatocyte growth-factor activator protein in turkey
RL (Meleagris gallopavo) deferent duct epithelial cells."
RN [2]
RP Comp. Biochem. Physiol. 132:769-777(2002).
SEQUENCE FROM N.A.
RA Holsberger D.R., Becker A.E., Thurston R.J., Rice C.D.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY216598; A046038.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR00742; EGF_2.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
PFam; PF00009; EGF_2.
PFam; PF00051; Kringle; 1.
PFam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
```

```
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FNL; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
FT NON TER 1
FT NON TER 540 540
SQ SEQUENCE 540 AA; 61159 MW; 0BB3B4A89C0B577F CRC64;

Query Match 31.8%; Score 252.5; DB 13; Length 540;
Best Local Similarity 39.5%; Pred. No. 4.3e-19;
Matches 49; Conservative 15; Mismatches 51; Indels 9; Gaps 2

QY 13 CLNGGTC---VSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRGKASTDTMG 68
Db 143 CMNGGCKMTASSGKTV-----CDCKGPFVGVKYNIVPNHHCYRGNGTEYRGTAKTISG 197
QY 69 RPLCPWNASATVLOQTY-HAHRSDALQLGLGKHNYCRNPNRRRPPRCYQVGLKPLVQECMV 128
Db 198 HSLCPWNSSLLYRELHVDVSVEKAVQLGLGPPSCYCRNPDEDEKPCWYIMKDNSLWSYECNI 257
QY 129 HDCA 132
Db 258 TSQA 261

RESULT 16
Q8VCS4 PRELIMINARY; PRT; 653 AA.
AC Q8VCS4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RC TISSUES=Liver;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC019376; AAH19376.1; -.
DR HSSP; P00761; IAN1.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR00742; EGF_2.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
PFam; PF00008; EGF; 2.
PFam; PF00039; fn1; 1.
PFam; PF00040; fn2; 1.
```

```

DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00018; EGF; 2.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00021; KRINGLE_2; 1.
DR PROSITE; PS00070; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
KW Kringle; Protease; Serine protease.
SQ SEQUENCE 653 AA; 70553 MW; FE18D90174ED6FDD CRC64;

Query Match 31.6%; Score 250.5; DB 11; Length 653;
Best Local Similarity 42.0%; Pred. No. 8.8e-19;
Matches 50; Conservative 11; Mismatches 41; Indels 17; Gaps 2;

QY 5 HQVPSNCCLNGGTCVSNKYFSNIHW-----CNPKFQGQHCIDKSKTCYEGNGH 56
Dd 239 HTACLSSPCLNGGTC-----HLIVGTGTSVCTPLGAGRCFNIVPTEHCFNGT 289
QY 57 FYRGKASTDTMGRPCLPWNSATVLQOYHAHRSALQLGLGKKNYCENPNRRPWCYV 115
Dd 290 EYRGVASTAAGSLSLANWSDLLIQELHVDVSAARVLLGLGPHAYCRNPKDERPWCYV 348

RESULT 17
O35727
ID O35727 PRELIMINARY; PRT; 597 AA.
AC O35727
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Factor XII.
GN F12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RA TISSUE=Liver;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; X9571; CAA67891.1; -.
DR HSRP; P00760; 1AQ7.
DR MEROPS; S01.211; -.
DR MGD; MGI:1891012; F12.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004235; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000083; FibrinctnI.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR008210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.

```

---

```

DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; FN1; 1.
DR Pfam; PF00040; FN2; 1.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 597 AA; 65638 MW; F3AC07C37D0C0FBA CRC64;

Query Match 31.4%; Score 249; DB 11; Length 597;
Best Local Similarity 40.8%; Pred. No. 1.2e-18;
Matches 49; Conservative 14; Mismatches 49; Indels 8; Gaps 3

QY 13 CLNGGTCVSNKYFSNIHWNCNPKFQGQHCIDKSKTCYEGNGHYPYRGKASTDTMGRPCL 72
Dd 183 CLNGGSC---LVEDHPLCRCPCTGYTFCDDLWATCVGRLGYRGAGTQTSGAPQ 239
QY 73 PWSNATVLQOY--HAHRSALQLGLGKKNYCENPNRRPWCYVQVGLKPLVQECNVHDC 131
Dd 240 RW-----TVEATYRNTEKQALSWGLQHAFCRNPNDRPWCVWSGDRLSWDYCGLEOC 295

RESULT 18
Q80YCS
ID Q80YCS PRELIMINARY; PRT; 609 AA.
AC Q80YCS;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to coagulation factor XII (Hageman factor) (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RA STRAIN=FVB/N; TISSUE=Liver;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC049867; AAH49867.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001891; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000083; FibrinctnI.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.

```

```
DR InterPro: IPR001314; Peptidase_S1A.
DR Pfam: PF00008; EGF_2.
DR Pfam: PF00039; fn1; 1.
DR Pfam: PF00040; fn2; 1.
DR Pfam: PF00051; kringie; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00013; FNTYPEII.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000995; FN_Type_II; 1.
DR ProDom: PD000395; Kringie; 1.
DR SMART: SM00181; EGF_2.
DR SMART: SM00179; EGF_CA; 2.
DR SMART: SM00058; FN1; 1.
DR SMART: SM00059; FN2; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; Tryp_SPC; 1.
DR PROSITE: PS00022; EGF_1; 2.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS01253; FIBRONECTIN_1; 1.
DR PROSITE: PS00023; FIBRONECTIN_2; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS00070; KRINGLE_2; 1.
DR PROSITE: PS0240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
DR NON_TER 1
SQ SEQUENCE 609 AA; 56783 MW; DF97D4DB236956D2 CRC64;

Query Match 31.4%; Score 249; DB 11; Length 609;
Best Local Similarity 40.8%; Pred. No. 1.2e-18;
Matches 49; Conservative 14; Mismatches 49; Indels 8; Gaps 3;

QY 13 CLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRGKASTDTWGRPCL 72
Db 195 CLNGGSL---LVEDHPICRCPTGYTCFDLDLWATCYEGSLSYRGAGTTSQAPCQ 251
QY 73 PWSATVLQOY--HAHRSDALQGLGKHNCRNPNRRRPMCYVQVGLKPLVQECMVHDC 131
Db 252 RW---TVBATYNNTEKQALSWGLGHAFCEPNPNDRPNCVNSGDRLSMDYCGLEQC 307

RESULT 19
Q81Z25 PRELIMINARY; PRT; 615 AA.
AC Q81Z25;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-WAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Coagulation factor XII-Mie.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wada H., Nishio K., Nakatani K., Kasai Y., Abe Y., Nobori T.;
RT Molecular characterization of coagulation factor XII-Mie.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB095845; BAC23095.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR009003; Cys Ser trypsin.
DR InterPro: IPR001881; EGF_CA.
DR InterPro: IPR006209; Fibrinctn1.
DR InterPro: IPR000083; Fibrinctn1.
DR InterPro: IPR000562; FN_Type_II.
DR InterPro: IPR006210; IEGF.
DR InterPro: IPR000001; Kringie.
DR InterPro: IPR001254; Peptidase_S1.
```

```
DR InterPro: IPR001314; Peptidase_S1A.
DR Pfam: PF00008; EGF_2.
DR Pfam: PF00039; fn1; 1.
DR Pfam: PF00040; fn2; 1.
DR Pfam: PF00051; kringie; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00013; FNTYPEII.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000995; FN_Type_II; 1.
DR ProDom: PD000395; Kringie; 1.
DR SMART: SM00181; EGF_2.
DR SMART: SM00179; EGF_CA; 2.
DR SMART: SM00058; FN1; 1.
DR SMART: SM00059; FN2; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; Tryp_SPC; 1.
DR PROSITE: PS00022; EGF_1; 2.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS01253; FIBRONECTIN_1; 1.
DR PROSITE: PS00023; FIBRONECTIN_2; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS00070; KRINGLE_2; 1.
DR PROSITE: PS0240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
DR NON_TER 1
SQ SEQUENCE 615 AA; 67735 MW; 030508870A0C7EDB CRC64;

Query Match 29.4%; Score 233; DB 4; Length 615;
Best Local Similarity 42.3%; Pred. No. 6.7e-17;
Matches 44; Conservative 12; Mismatches 40; Indels 8; Gaps 3;

QY 13 CLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRGKASTDTWGRPCL 72
Db 183 CLHGGCLB---VEGRLCHCPVGTGPFCDVDTASCYDGRGLYRLARTILSGAPCQ 239
QY 73 PWSATVLQOY--HAHRSDALQGLGKHNCRNPNRRRPMCYV 115
Db 240 PWSA-----EATYRNVTAEQARNWGLGHAFCEPNPNDRPNCV 279

RESULT 20
Q9BZW1 PRELIMINARY; PRT; 395 AA.
AC Q9BZW1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Neonatal thrombolytic agent alpha-form (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Dou D.;
RT "A brain-type plasminogen activator.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF260825; AAK11956.1; -.
DR HSP; P00750; 1PK2.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR009003; Cys Ser trypsin.
DR InterPro: IPR000083; Fibrinctn1.
DR InterPro: IPR000001; Kringie.
DR InterPro: IPR001254; Peptidase_S1.
DR InterPro: IPR001314; Peptidase_S1A.
DR Pfam: PF00039; fn1; 1.
```

DR Pfam: PF00051; kringle; 1.  
DR Pfam: PF00089; trypsin; 1.  
DR PRINTS: PR00722; KRWCRYPIN.  
DR PRINTS: PR00018; KRINGLE.  
DR ProDom: PD000395; Kringle; 1.  
DR SMART: SM00058; FN1; 1.  
DR SMART: SM00130; KR; 1.  
DR SMART: SM00020; Tryp Spc; 1.  
DR PROSITE: PS01253; FIBRONECTIN\_1; 1.  
DR PROSITE: PS00021; KRINGLE\_1; 1.  
DR PROSITE: PS00070; KRINGLE\_2; 1.  
DR PROSITE: PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE: PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.  
FT NON\_TER 395 395  
SQ SEQUENCE 395 AA; 44323 MW; 3FBD4A2F0B7C11C8 CRC64;  
  
Query Match 28.9%; Score 229.5; DB 4; Length 395;  
Best Local Similarity 42.7%; Pred. No. 1e-16;  
Matches 44; Conservative 10; Mismatches 44; Indels 5; Gaps 2;  
  
QY 30 WNCNPKFGGHCETDKSKCYEGNGHYGKASDTWTGRCLPWSATVLQOYTHAHS 89  
DB 33 WCNSS----GRAQCS-EGNSDCYFGNSAYRGTHSLTESGASCLPWSMILIGKVTYQNP 87  
  
QY 90 DALQLGLGKHNYCRNPNRRPWCYVQVGLKPLVQECMVHDC 132  
DB 88 SAQALGLGKHNYCRNPNRPGDKAPWCHVKNRRLTWEYCDVPSCS 130  
  
RESULT 21  
Q8NG20 PRELIMINARY; PRT; 90 AA.  
AC Q8NG20;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Plasmogen/activator kringle.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Dou D.;  
RT "Production of kringle fragment."  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  
DR EMBL: AF282882; AAM5248.1; -.  
DR Pfam: PF00051; kringle; 1.  
DR PRINTS: PR00018; Kringle; 1.  
DR ProDom: PD000395; Kringle; 1.  
DR SMART: SM00130; KR; 1.  
DR PROSITE: PS00021; KRINGLE\_1; 1.  
DR PROSITE: PS00070; KRINGLE\_2; 1.  
KW Glycoprotein; Kringle.  
SQ SEQUENCE 90 AA; 9504 MW; A33887F9FD4C7B1 CRC64;  
  
Query Match 27.6%; Score 218.5; DB 4; Length 90;  
Best Local Similarity 49.4%; Pred. No. 3.2e-16;  
Matches 41; Conservative 5; Mismatches 36; Indels 1; Gaps 1;  
  
QY 50 CYEGNGHYGKASDTWTGRCLPWSATVLQOYTHAHSRALQLGLGKHNYCRNPNR 109  
DB 8 CYFGNSAYRGTHSLTESGASCLPWSMILIGKVTYQNPQAQALGLGKHNYCRNPDGA 67  
  
QY 110 RPWCYVQVGLKPLVQECMVHDC 132  
DB 68 KWCYNT-TNPRKLYCDVPQCA 89

RESULT 22  
Q90675 PRELIMINARY; PRT; 202 AA.  
ID Q90675;  
AC Q90675;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Tissue-type plasminogen activator (Fragment).  
GN TPA.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=White Leghorn;  
RX MEDLINE=97199025; PubMed=9047000;  
RT Johnson A.L., Bridgham J.T., Anthony R.V.;  
RT "Expression of avian urokinase and tissue-type plasminogen activator  
RT messenger ribonucleic acid during follicle development and atresia.";  
RL Biol. Reprod. 56:581-588(1997).  
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.  
DR EMBL: U31986; AAA74955.1; -.  
DR HSP; P00750; IRTF.  
DR MEROPS; S01.232; -.  
DR GO: GO:0008233; F:peptidase activity; IEA.  
DR GO: GO:0004295; F:trypsin activity; IEA.  
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro: IPR009003; Cys\_Ser\_trypsin.  
DR InterPro: IPR000001; Kringle.  
DR InterPro: IPR001254; Peptidase\_S1.  
DR Pfam: PF00051; kringle; 1.  
DR Pfam: PF00089; trypsin; 1.  
DR PRINTS: PR00018; Kringle.  
DR ProDom: PD000395; Kringle; 2.  
DR SMART: SM00130; KR; 1.  
DR PROSITE: PS00021; KRINGLE\_1; 1.  
DR PROSITE: PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE: PS00134; TRYPSIN\_HIS; 1.  
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.  
FT NON\_TER 1 1  
FT NON\_TER 202 202  
SQ SEQUENCE 202 AA; 23148 MW; 049DB42941D83AFC CRC64;  
  
Query Match 24.9%; Score 197.5; DB 13; Length 202;  
Best Local Similarity 36.9%; Pred. No. 1.5e-13;  
Matches 41; Conservative 11; Mismatches 50; Indels 9; Gaps 2  
  
QY 30 WNCNPK--KEGQHCE-----IDSKTCYEGNGHYGKASDTWTGRCLPWSATVL 80  
DB 12 WCYVFKAGYIIEFCSTPACTKVAEDGDCYTGNGLAYRGTRSKSGFSCLPWNPVFLT 71  
  
QY 81 QQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLKPLVQECMVHDC 131  
DB 72 SKIYTALESQRRALGLGKHNYCRNPNRPGDAQPNCHVWKDRLTWEYCDVPQC 122  
  
RESULT 23  
Q8AXX3 PRELIMINARY; PRT; 421 AA.  
ID Q8AXX3;  
AC Q8AXX3;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Kremen2.  
GN KEM2.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xeropodinae; Xenopus.  
OX NCBI\_TaxID=8355;



```

DR GO:0003809; F:thrombin activity; IEA.
DR GO:0004295; F:trypsin activity; IEA.
DR GO:0007536; P:blood coagulation; IEA.
DR GO:0006388; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR009003; Cys_Ser_trypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003609; Pan_app.
DR InterPro: IPR001254; Peptidase_S1.
DR InterPro: IPR001314; Peptidase_S1A.
DR InterPro: IPR003966; Peptidase_S1A_pr.
DR Pfam: PF00051; kringle; 4.
DR Pfam: PF00024; PAN; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR PRINTS: PR01505; PROTHROMBIN.
DR ProDom: PD000395; Kringle; 4.
DR SMART: SM00130; KR; 4.
DR SMART: SM00473; PAN_AP; 1.
DR SMART: SM00020; TRY_SPC; 1.
DR PROSITE: PS00021; KRINGLE_1; 4.
DR PROSITE: PS50070; KRINGLE_2; 4.
DR PROSITE: PS0240; TRYPSIN_DOM; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
SQ SEQUENCE 716 AA; 81971 MW; 508376A0E4398798 CRC64;

Query Match 20.9%; Score 166; DB 13; Length 716;
Best Local Similarity 31.1%; Pred. No. 1.6e-09;
Matches 42; Conservative 11; Mismatches 40; Indels 42; Gaps 6;

QY 30 WC-----NCPKFGQGHCEIDSK-----TCVEGNHGYRGKASTDTMGRCLP 73
Db 258 WCYTDPNVEKEF-----CRITCKQRQLSNIETSTCFKEGEGYRKANTTSGIPCQR 313
QY 74 WNSAT-----VLQOTYHAHRSDALQLGLGKHNCRPNDRRRPWCYVQV-----GL 119
Db 314 WDSQTFQSHRFLPEKYPCKGLD-----ENYCRNPDGSEAPCFTLLPGRMAYCFQI 365
QY 120 KPLVQECMVHDCADG 134
Db 366 KRCKDDVLEPDCYHG 380

RESULT 26
Q8WNR1 PRELIMINARY; PRT; 359 AA.
AC Q8WNR1;
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Plasmicogen (Fragment).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Pirie-Shepherd S.R., Coffman K.T., Resnick D., Chan R., Kisker O.,
RA Foltman J., Waters D.J.,
RA "Angiostatin is detectable in the Urine of Dogs with Spontaneous Bone
RT Cancer.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
DR EMBL; AY069985; AAL8519.1; -.
DR GO:0005509; F:calcium ion binding; IEA.
DR GO:0003803; F:thrombin activity; IEA.
DR GO:0007596; F:blood coagulation; IEA.
DR GO:0006508; F:proteolysis and peptidolysis; IEA.
DR InterPro: IPR000001; Kringle.
DR Pfam: PF00051; kringle; 4.
DR PRINTS: PR00018; KRINGLE.

```

```

DR PRINTS: PR01505; PROTHROMBIN.
DR ProDom: PD000395; Kringle; 4.
DR SMART: SM00130; KR; 4.
DR PROSITE: PS00021; KRINGLE_1; 4.
DR PROSITE: PS50070; KRINGLE_2; 4.
KW Glycoprotein; Kringle.
FT NON TER 1
FT NON TER 359
SQ SEQUENCE 359 AA; 41172 MW; 776D35F4AB0BDD9E CRC64;

Query Match 20.9%; Score 165.5; DB 6; Length 359;
Best Local Similarity 29.0%; Pred. No. 8.7e-10;
Matches 42; Conservative 13; Mismatches 43; Indels 47; Gaps 7

QY 5 HQVPSNCDCLNGTCTVSNKYFSN-----IHWG-NCPKFGQGHCEIDSKT----- 49
Db 211 NRTPEFPCKN-----LDENYCRNPDGSEAPWCYTTNSEVRWEHCQIPCESSPITTEYLD 266
QY 50 -----CYEGNHGYRGKASTDTMGRCLPWN SATVLQOTYHAHRSDAL-- 92
Db 267 APASVPPEQTPVQECYHGSGYRGTSSTTTTGRKCSWSMT-----PHRHKTP 319
QY 93 ---OLGLGKHNCRPNDRRRPWCY 114
Db 320 HPFEAGL-TMNYCRNPDADKSPWCY 343

RESULT 27
Q9ROW3 PRELIMINARY; PRT; 812 AA.
AC Q9ROW3;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Plasmicogen protein precursor (EC 3.4.21.7).
GN PLASMINOGEN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Bangert K., Johnsen A.H., Thorsen S.;
RT "Rat plasminogen: cDNA and gene structure.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=91350378; PubMed=1645711;
RA Kanalas J.J., Makker S.P.;
RT "Identification of the rat Heymann nephritis autoantigen (GP330) as a
RT receptor site for plasminogen.";
RL J. Biol. Chem. 266:10825-10829(1991).
CC -!- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
DR EMBL; AJ242649; CAB46014.1; -.
DR HSSP; P00747; 1PMK.
DR GO:0005576; C:extracellular; IEA.
DR GO:0005509; F:calcium ion binding; IEA.
DR GO:0004263; F:chymotrypsin activity; IEA.
DR GO:0005179; F:hormone activity; IEA.
DR GO:0008233; F:peptidase activity; IEA.
DR GO:0004283; F:plasmin activity; IEA.
DR GO:0003809; F:thrombin activity; IEA.
DR GO:0004295; F:trypsin activity; IEA.
DR GO:0007596; F:blood coagulation; IEA.
DR GO:0006508; F:proteolysis and peptidolysis; IEA.
DR InterPro: IPR009003; Cys_Ser_trypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003609; Pan_app.
DR InterPro: IPR001254; Peptidase_S1.
DR InterPro: IPR001314; Peptidase_S1A.

```

```
DR InterPro; IPR003966; Peptidase_S1A_pr.
DR InterPro; IPR001400; Somatotropin.
DR Pfam; PF00051; kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE_1; 1.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN AP; 1.
DR SMART; SM00020; TYD_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 5.
DR PROSITE; PS00070; KRINGLE_2; 5.
DR PROSITE; PS00338; SOMATOTROPIN_2; 1.
DR PROSITE; PS02240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease; Signal.
FT SIGNAL 1 19
FT CHAIN 20 812 PLASMINOGEN.
SQ SEQUENCE 812 AA; 90535 MW; 8C703C51410EB9C9E CRC64;

Query Match 20.8%; Score 165; DB 11; Length 812;
Best Local Similarity 29.0%; Pred. No. 2.4e-09;
Matches 47; Conservative 20; Mismatches 49; Indels 46; Gaps 11;

QY 5 HOVPSNCDCLNGGTCVSNKYFSN-----IHWC-NCPKXFGGQHCEI-----DKS-- 47
DB 308 NRTPENPPCKN-----LEENTCRPPDGETAPWCYTTDSQLRWEYCEIPSCGSSVSPQSDS 363
QY 48 -----KTCYEGNGHYFGKASTDTMGPCLPFNASATVLCQTYHAHRSDALQL--- 94
DB 364 SVLPQETPVVQECVQNGKSYRGTSITNTGKKQSW----VSMTPHSHSXTPANFPDA 418
QY 95 GLGKHYCRPNDR-RRPWCYVQVGLKPLY--QECMVHDCAD 133
DB 419 GL-EMNYCRPNDRQDGPWCFT---TDPVWRWYCNLKRCS 456

RESULT 28
Q9BRB6 PRELIMINARY; PRT; 393 AA.
ID Q9BRB6
AC Q9BRB6
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to receptor tyrosine kinase-like orphan receptor 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC006374; AA006374.1; -.
DR HSP; P00747; ICEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0004988; F:transmembrane receptor activity; IEA.
DR GO; GO:0007275; P:development; IEA.
DR InterPro; IPR000024; Fz_domain.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF00047; Ig; 1.
DR PRINTS; PR00051; kringle; 1.
DR ProDom; PD000395; Kringle; 1.

QY 5 HOVPSNCDCLNGGTCVSNKYFSN-----IHWC-NCPKXFGGQHCEI-----DKS-- 47
DB 308 NRTPENPPCKN-----LEENTCRPPDGETAPWCYTTDSQLRWEYCEIPSCGSSVSPQSDS 363
QY 48 -----KTCYEGNGHYFGKASTDTMGPCLPFNASATVLCQTYHAHRSDALQL--- 94
DB 364 SVLPQETPVVQECVQNGKSYRGTSITNTGKKQSW----VSMTPHSHSXTPANFPDA 418
QY 95 GLGKHYCRPNDR-RRPWCYVQVGLKPLY--QECMVHDCAD 133
DB 419 GL-EMNYCRPNDRQDGPWCFT---TDPVWRWYCNLKRCS 456

RESULT 28
Q9BRB6 PRELIMINARY; PRT; 393 AA.
ID Q9BRB6
AC Q9BRB6
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to receptor tyrosine kinase-like orphan receptor 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC006374; AA006374.1; -.
DR HSP; P00747; ICEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0004988; F:transmembrane receptor activity; IEA.
DR GO; GO:0007275; P:development; IEA.
DR InterPro; IPR000024; Fz_domain.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF00047; Ig; 1.
DR PRINTS; PR00051; kringle; 1.
DR ProDom; PD000395; Kringle; 1.
```

```
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00038; FZ; 1.
DR PROSITE; PS00835; IG LIKE; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Immunoglobulin domain; Kinase; Kringle; Receptor.
SQ SEQUENCE 393 AA; 43825 MW; 1F93DCBBBF53855 CRC64;

Query Match 20.0%; Score 158.5; DB 4; Length 393;
Best Local Similarity 29.7%; Pred. No. 5.5e-09;
Matches 43; Conservative 14; Mismatches 51; Indels 37; Gaps 7

QY 2 NELHQVP-----SNCDCLNGGTCVSNKYFS-----NIHWNC-----PKFGGQH 41
DB 236 DETSSVPKPRDLCDRDECEILENVLCOTEYIFARSNPMLMLRLKLPNCEDLPQESPEAAN 295
QY 42 C-----BIDSKTCYEGNGHYFGKASTDTMGPCLPFNASATVLCQTYHAHRSDAL 92
DB 296 CIRIGIPWADPINKNKKCYNSTGVDYRGTVSVTKSGRQCQFWS-----QYPHTHTFAL 350
QY 93 QLG--LGKHYCRPNDRRR-PMCY 114
DB 351 RPELNGGHSYCRPNQKQKAPWCF 375

RESULT 29
Q9BNP9 PRELIMINARY; PRT; 937 AA.
ID Q9BNP9
AC Q9BNP9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Receptor tyrosine kinase-like orphan receptor 1.
CN ROR1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Adipose tissue;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK080896; BAC38068.1; -.
DR MGD; MGI:1347520; Ror1.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0004888; F:transmembrane receptor activity; IEA.
DR GO; GO:0007275; P:development; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000024; Fz_domain.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00069; kinase; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000395; Kringle; 1.
```

```
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00220; S.TKC; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00038; FZ; 1.
DR PROSITE; PS00335; IG LIKE; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00125; TYR PKINASE.
SQ SEQUENCE 937 AA; 104118 MW; CB440323CA66EC28 CRC64;

Query Match
Best Local Similarity 20.0%; Score 158.5; DB 11; Length 937;
Matches 43; Conservative 14; Mismatches 51; Indels 37; Gaps 7;

QY 2 NELHQP-----SNCDCLNGGTCVSNKYFS-----NIHWNC-----PKKFGGQH 41
DB 236 DETSSVPRDLRCDECEVLENVLCQTEYIFARSNPMLMLRLKLPNCEDLPQSPESPAAN 295
QY 42 C-----EIDSKTCTVEGNGHFYRGKASDTMGRCPLPNWSATVLCQTYHAHRSAL 92
DB 296 CIRIGIPMDPINKHKCNSTGVDRGTVSVTKSGROCPWNS-----QYPHTHSFTAL 350
QY 93 QLG--LGKHNCRNPNRRR--PCWY 114
DB 351 RPEELNGHSHYCRNPGNKEAPWCF 375

RESULT 30
QSBG10 PRELIMINARY; PRT; 937 AA.
ID QSBG10
AC QSBG10;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Receptor tyrosine kinase-like orphan receptor 1.
GN ROR1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Adipose tissue;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573 (2002).";
DR EMBL; AK046699; BAC32840.1; -.
DR MGD; MGI:1347520; Ror1.
DR GO; GO:0016020; C.membrane; IEA.
DR GO; GO:0005524; P.ATP binding; IEA.
DR GO; GO:0004674; P.protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; P.protein-tyrosine kinase activity; IEA.
DR GO; GO:0004888; P.transmembrane receptor activity; IEA.
DR GO; GO:0007275; P.development; IEA.
DR GO; GO:0006468; P.protein amino acid phosphorylation; IEA.
DR InterPro; IPR000024; Fz domain.
DR InterPro; IPR003599; Ig-like.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR InterPro; IPR008266; Tyr_pkinase_AS.
PFam; PF01392; Fz; 1.

Query Match
Best Local Similarity 20.0%; Score 158.5; DB 11; Length 937;
Matches 43; Conservative 14; Mismatches 51; Indels 37; Gaps 7;

QY 2 NELHQP-----SNCDCLNGGTCVSNKYFS-----NIHWNC-----PKKFGGQH 41
DB 236 DETSSVPRDLRCDECEVLENVLCQTEYIFARSNPMLMLRLKLPNCEDLPQSPESPAAN 295
QY 42 C-----EIDSKTCTVEGNGHFYRGKASDTMGRCPLPNWSATVLCQTYHAHRSAL 92
DB 296 CIRIGIPMDPINKHKCNSTGVDRGTVSVTKSGROCPWNS-----QYPHTHSFTAL 350
QY 93 QLG--LGKHNCRNPNRRR--PCWY 114
DB 351 RPEELNGHSHYCRNPGNKEAPWCF 375

RESULT 31
Q46506 PRELIMINARY; PRT; 454 AA.
ID Q46506
AC Q46506;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Apolipoprotein a (Fragment).
GN BAAPOA.
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Papio.
OX NCBI_TaxID=9557;
RN [1]
RP SEQUENCE FROM N.A.
RA Cox L.A., Jett C., Hixson J.E.;
RT "Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice
RT Site Mutation is Associated with Deletion of a Single Exon in a Null
RT Allele.";
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AF029691; AAB97886.1; -.
DR HSSP; P00747; 2PK4.
DR MEROPS; S01.999; -.
DR GO; GO:0004263; F.chymotrypsin activity; IEA.
DR GO; GO:0008233; F.peptidase activity; IEA.
DR GO; GO:0004295; F.trypsin activity; IEA.
DR GO; GO:0006508; P.proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_ser_trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
PFam; PF00031; kringle; 2.
PFam; PF00089; trypsin; 1.
```



Db 64 DADTGPWCFT--MDPSIRWEYCNLTRCSD 90

RESULT 33

Q7TP84 PRELIMINARY; PRT; 759 AA.

ID	Q7TP84	AC	Q7TP84	DT	01-OCT-2003 (TReMBLrel. 25, Created)	DT	01-OCT-2003 (TReMBLrel. 25, Last sequence update)	DT	01-OCT-2003 (TReMBLrel. 25, Last annotation update)	DE	Abi-346.
OS	Rattus norvegicus (Rat).	OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	OC	NCBI_TaxID=10116;	OC	NCBI_TaxID=10116;	RN	[1]
RP	SEQUENCE FROM N.A.	RA	Xu C.S., Li W.Q., Li Y.C., Yan H.M., Chang C.F., Zhao L.F., Ma H.,	RA	Wang L., Wang S.F., Han H.P., Wang G.P., Chai L.Q., Yuan J.Y.,	RA	Yang K.J., Shi J.B., Rahman S., Wang Q.N., Zhang J.B.;	RT	"Liver regeneration after PH";	RL	Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AY325159; AAP92560.1.	DR	EMBL; AY325159; AAP92560.1.	DR	EMBL; AY325159; AAP92560.1.	DR	EMBL; AY325159; AAP92560.1.	DR	EMBL; AY325159; AAP92560.1.	DR	EMBL; AY325159; AAP92560.1.
SQ	SEQUENCE 759 AA; 86056 MW; 1422BFAC05C6DFA7 CRC64;	SQ	SEQUENCE 759 AA; 86056 MW; 1422BFAC05C6DFA7 CRC64;	SQ	SEQUENCE 759 AA; 86056 MW; 1422BFAC05C6DFA7 CRC64;	SQ	SEQUENCE 759 AA; 86056 MW; 1422BFAC05C6DFA7 CRC64;	SQ	SEQUENCE 759 AA; 86056 MW; 1422BFAC05C6DFA7 CRC64;	SQ	SEQUENCE 759 AA; 86056 MW; 1422BFAC05C6DFA7 CRC64;

Query Match 19.7%; Score 156; DB 11; Length 759;  
 Best Local Similarity 24.9%; Pred. No. 2.1e-08;  
 Matches 50; Conservative 20; Mismatches 47; Indels 84; Gaps 11

QY 5 HOVPSNCDCNGGTCVSNKYFSN-----IHW-CNPKKFGQHCEI-----DKS-- 47

Db 316 NRTPEPFCKK-----LEENYCRNPDGTAPWCYTTSQLRWEYCEIPSGSSVSPDQSDS 371

QY 48 -----KTCYEGNGHYRGKASTDTMGRCPLPWSATVLQOTYHAHS-----DA 91

Db 372 SVLPQPTFWQECYQGNKSYRGTSITNTGKQCSW-----VSMTPHSHKSTPANFPDA 426

QY 92 LQLGLGKH-----NYCRPDPN-RRPWCY 114

Db 427 YONLIQIQRSVSYVINGSLWETFLIHTNQKYSVQLKESGLEMMNYCRPNDQDQGPWC 486

QY 115 VQVGLKPLV--QECWVHDCAD 133

Db 487 T--TDPVSRYEYCNLKRCE 504

RESULT 34

Q7SXB3 PRELIMINARY; PRT; 263 AA.

ID	Q7SXB3	AC	Q7SXB3	DT	01-OCT-2003 (TReMBLrel. 25, Created)	DT	01-OCT-2003 (TReMBLrel. 25, Last sequence update)	DT	01-OCT-2003 (TReMBLrel. 25, Last annotation update)	DE	Hypothetical protein.		
OS	Brachydanio rerio (Zebrafish) (danio rerio).	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	OC	Actinopterygii; Neopterygii; Teleostei; Osteiophysi; Cypriniformes;	OC	Cyprinidae; Danio.	OC	NCBI_TaxID=7955;	RN	[1]		
RP	SEQUENCE FROM N.A.	RA	STRAIN=AB; TISSUE=Body;	RA	MEDLINE=22368257; PubMed=12477932;	RA	Klausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,	RA	Klausberg R.D., Collins F.S., Wagner L., Shennen C.M., Schuler G.D.,	RA	Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,	RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,	RA	Stapleton M., Soares M.B., Boraldo M.F., Casavant T.L., Scheetz T.E.,	RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Cacciari P., Prange C.,	RA	Raha S.S., McQuillan N.A., Peters G.J., Abramson R.D., Mullaly S.J.,	RA	Rosak S.A., McEwan P.J., Malek J.A., Gunaratne P.H.	RA	SEQUENCE FROM N.A.

Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J., Marra M.A.;  
"genome and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";  
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[2]  
SEQUENCE FROM N.A.  
STRAIN=AB; TISSUE=Body;  
Strausberg R.;  
Submitted (Aug-2003) to the EMBL/GenBank/DBJ databases.  
EMBL; BC055675; AAH55675.1; -  
Hypothetical protein.  
KW  
SEQUENCE 263 AA; 28777 MW; 8BEBG117EC7C8A58 CRC64;  
Query Match 19.6%; Score 155.5; DB 13; Length 263;  
Best Local Similarity 35.6%; Pred. No. 7.6e-09;  
Matches 31; Conservative 8; Mismatches 39; Indels 9; Gaps 2;  
QY 48 KTCYEGNGHFYRGKASTDTMGRLCPNWSATVLQQTYYHAHRSDALQLGLGKINYCRNPDN 107  
Db 23 KDCITNNGEDYRGYQKTKSSGTCLSRSLNL-----KFKDSGTGVGDHFCRNPDG 74  
QY 108 RRRPWCYQVGLKPLVQE-CMVHDCAD 133  
Db 75 SNKPMCYVSGSGGTKEACDIRICQD 101  
RESULT 35  
O18783  
ID Q18783 PRELIMINARY; PRT; 806 AA.  
AC Q18783;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Plasminogen.  
OS Macropus eugenii (Tamar wallaby).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.  
OC NCBI\_TaxID=9315;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=98004511; PubMed=9342350;  
RA Lawn R.M., Schwartz K., Pathy L.;  
RT "Convergent evolution of apolipoprotein(a) in primates and hedgehog.";  
RT Proc. Natl. Acad. Sci. U.S.A. 94:11992-11997(1997).  
CC -I- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.  
DR EMBL; AF012297; AAB65760.1; -  
DR HSSP; P00747; 5HPG.  
DR MEROPS; S01.233; -  
DR GO; GO:0005509; F.calcium ion binding; IEA.  
DR GO; GO:0004263; F.chymotrypsin activity; IEA.  
DR GO; GO:0008233; F.peptidase activity; IEA.  
DR GO; GO:0003809; F.thrombin activity; IEA.  
DR GO; GO:0004295; F.trypsin activity; IEA.  
DR GO; GO:0007596; P.blood coagulation; IEA.  
DR GO; GO:0006508; P.proteolysis and peptidolysis; IEA.  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR003014; PAN.  
DR InterPro; IPR003609; Pan\_app.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR InterPro; IPR003966; Peptidase\_S1A\_pr.  
DR Pfam; PF00051; kringle; 5.  
DR Pfam; PF00024; PAN; 1.  
DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PRO0018; KRINGLE.  
DR PRODOM; PRO1505; PROTHROMBIN.  
DR PRODOM; PD000395; Kringle; 5.  
DR SMART; SM00130; KR; 5.  
DR SMART; SM00473; PAN AP; 1.  
DR SMART; SM00020; TRYPSIN; 1.  
DR PROSITE; PS00021; KRINGLE 1; 5.  
DR PROSITE; PS00070; KRINGLE 2; 5.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
DR Glycoprotein; Hydrolase; Kringle; Protease; Serine protease; Signal.  
FT SIGNAL 1 19 POTENTIAL.  
FT CHAIN 20 810 PLASMINOGEN.  
SQ SEQUENCE 810 AA; 90555 MW; B05C7D4B0D020B3C CRC64;  
  
Query Match 19.4%; Score 154; DB 4; Length 810;  
Best Local Similarity 28.0%; Pred. No. 3.8e-08;  
Matches 45; Conservative 20; Mismatches 50; Indels 46; Gaps 10;  
  
QY 5 HQVPSNCDCLNGTGVSNKYFN-----IHWNCNP-KKFGGQHCEI-----DKS-----47  
Db 308 NRTPEFPCKN---LDENYCNPDGKRAPWCHTNSQVRWEYCKIPSCDSSPVSTEQLA 363  
  
QY 48 -----KTCYEGNGHFRGKASDTMGRCPLPWSATVLOQTYHAHR-----SDALQ 93  
Db 364 PTAPPELTVPVQDCHGDSYRGTSTTTGKQCSWSS-----MTPRHQKTPENYPN 418  
  
QY 94 LGLGKHNYCRNPNRRPVCYQVGLKPLV--QECWVHDC 132  
Db 419 AGL-TWNYCNPDADKGPFCFT---TDPVSVRWEYCNLKKCS 455  
  
RESULT 37  
Q9UIR7  
ID Q9UIR7 PRELIMINARY; PRT; 113 AA.  
AC Q9UIR7;  
DT 01-MAY-2000 (T-EMBLrel. 13, Created)  
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)  
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)  
DE Apolipoprotein(a) (Fragment).  
GN APOA.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21181705; PubMed=11285247;  
RA Ogorelkova M., Kraft H.G., Ehnholm C., Utermann G.;  
RT "Single nucleotide polymorphisms in exons of the apo(a) kringles IV  
RT types 6 to 10 domain affect lip(a) plasma concentrations and have  
RT different patterns in Africans and Caucasians."  
RL Hum. Mol. Genet. 10:815-824(2001).  
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  
DR EMBL; AF158659; AAF03678.1; -  
DR EMBL; AF158658; AAF03678.1; JOINED.  
DR HSSP; P00747; 2PK4.  
DR InterPro; IPR000001; Kringle.  
DR Pfam; PF00051; kringle; 1.  
DR PRINTS; PR00018; KRINGLE.  
DR PRODOM; PD000395; Kringle; 1.  
DR SMART; SM00130; KR; 1.  
DR PROSITE; PS00021; KRINGLE 1; 1.  
DR PROSITE; PS00070; KRINGLE 2; 1.  
DR Glycoprotein; Kringle; Lipoprotein.  
FT NON\_TER 1 1  
FT NON\_TER 113 113  
SQ SEQUENCE 113 AA; 12815 MW; 4F80ADF8708548CB CRC64;  
  
Query Match 19.3%; Score 153; DB 4; Length 113;  
Best Local Similarity 39.3%; Pred. No. 5.7e-09;  
Matches 33; Conservative 8; Mismatches 35; Indels 8; Gaps 3;

QY 50 CYEGNGHFRGKASDTMGRCPLPWSATVLOQTYHAHRSDALQGLGKHNYCRNPNRR 109  
Db 11 CYHGDGQSIRGSPSTVTTGRTQCSWSSMT---PHMQRTTEYYPNGGLTRNYCRNPDABI 67  
  
QY 110 RPYCYVQVGLKPLV--QECWVHDC 131  
Db 68 RPYCYT---MDPSVRWEYCNLTRC 88  
  
RESULT 38  
P70006  
ID P70006 PRELIMINARY; PRT; 717 AA.  
AC P70006;  
DT 01-FEB-1997 (T-EMBLrel. 02, Created)  
DT 01-FEB-1997 (T-EMBLrel. 02, Last sequence update)  
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
DE Hepatocyte growth factor-like protein precursor.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RX MEDLINE=96404125; PubMed=8808403;  
RA Aberger F., Schmidt G., Richter K.;  
RT "The Xenopus homologue of hepatocyte growth factor-like protein is  
RT specifically expressed in the presumptive neural plate during  
RT gastrulation."  
RL Mech. Dev. 54:23-37(1996).  
CC -1- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.  
DR EMBL; Y08734; CAA69989.1; -  
DR HSSP; P00747; ICEA.  
DR MEROPS; S01.977; -  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0003809; F:thrombin activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0007596; P:blood coagulation; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR009003; Cys Ser trypsin.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR003014; PAN.  
DR InterPro; IPR003609; Pan\_app.  
DR InterPro; IPR01254; Peptidase\_Si.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR InterPro; IPR003966; Peptidase\_S1A\_pr.  
DR Pfam; PF00051; kringle; 4.  
DR Pfam; PF00024; PAN; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHIMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR PRINTS; PR01505; PROTHROMBIN.  
DR PRODOM; PD000395; Kringle; 4.  
DR SMART; SM00130; KR; 4.  
DR SMART; SM00473; PAN AP; 1.  
DR SMART; SM00020; TRYPSIN; 1.  
DR PROSITE; PS00021; KRINGLE 1; 4.  
DR PROSITE; PS00070; KRINGLE 2; 4.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR Glycoprotein; Hydrolase; Kringle; Protease; Serine protease; Signal.  
FT SIGNAL 1 28 POTENTIAL.  
FT CHAIN 29 717 HEPATOCYTE GROWTH FACTOR-LIKE PROTEIN.  
SQ SEQUENCE 717 AA; 82017 MW; 6F877A432C8DD54 CRC64;  
  
Query Match 19.2%; Score 152.5; DB 13; Length 717;  
Best Local Similarity 31.7%; Pred. No. 4.8e-08;  
Matches 39; Conservative 13; Mismatches 42; Indels 29; Gaps 6  
  
QY 30 WC-----NCPKFGQHCEIDKSK-----TCYEGNGHFRGKASDTMGRCPLP 73

```

Db 259 WCYTDPNVEREF-----CHITCKEQRISNIEITSTCFKRGEGYRGKANTTTSIGPCQR 314
QY 74 WNSATVLOQTYHAHRSALQL---GLGKHNYCRNPDNRRPWCYVQVGLKPLVQECMVD 130
Db 315 WDT-----QAPHVHFLPKYCKGLDE-NYCRNPVGSAPWCFTTLKNRMWYCFQIKR 368
QY 131 CAD 133
Db 369 CTD 371

RESULT 39
QSH1V4
ID Q9H1V4 PRELIMINARY; PRT; 648 AA.
AC Q9H1V4;
DT 01-NAR-2001 (TrEMBLrel. 16, Created)
DT 01-NAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE DJ1182A14.3 (Similar to MSt1 (Macrophage stimulating 1 (Hepatocyte
growth factor-like)))
CN DJ1182A14.3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bird C.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
DR EMBL; AL137798; CAC17639.1; -.
DR HSSP; P00747; SHPG.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0008508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00051; kringle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0018; KRINGLE.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; TRYSP; 1.
DR PROSITE; PS00021; KRINGLE_1; 3.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
SQ SEQUENCE 648 AA; 72781 MW; 4C5077057350E463 CRC64;

Query Match 19.1%; Score 151.5; DB 4; Length 648;
Best Local Similarity 27.9%; Pred. No. 5.5e-08;
Matches 36; Conservative 18; Mismatches 46; Indels 29; Gaps 6;

QY 30 WC-NCPKFGGQHCEIDK-----SKTCYEGNGHFYRGKASTDTMGRPCLPWN 76
Db 205 WCYTDPQIEREFCDLPRCGSEAPQREATSVSCFRKGEGYRGANTTTAGVPCQRWDA 264
QY 77 ATVLQQTVAHRSALQLGLG---KHNYCRNPDNRRPWCY-----VOVGLKPLVQEC--- 126
Db 265 -----QIPHQRFTPEKYACKDLRENFCEPNPGSEAPWCFTLPCRMVGFYQIRCTDD 319
QY 127 -MWHDCADG 134

```

```

Db 320 VREQDCYHG 328
RESULT 40
Q13208
ID Q13208 PRELIMINARY; PRT; 567 AA.
AC Q13208;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hepatocyte growth factor-like protein homolog.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Degen S.J.F.; McDowell S.A.; Waltz S.E.; Gould F.; Stuart L.A.;
Carritt B.;
RL DNA Seq. 8:409-413(1998).
CC -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
DR EMBL; U28054; AAC63092.1; -.
DR HSSP; P00747; 2PK4.
DR MRO8S; S01.977; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0003809; F:thrombin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0007596; P:blood coagulation; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR003966; Peptidase_S1A_pr.
DR Pfam; PF00051; kringle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0018; KRINGLE.
DR PRINTS; PRO1505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00020; TRYSP; 1.
DR PROSITE; PS00021; KRINGLE_1; 3.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
SQ SEQUENCE 567 AA; 64117 MW; 3FC38B07F1645810 CRC64;

Query Match 19.0%; Score 150.5; DB 4; Length 567;
Best Local Similarity 27.9%; Pred. No. 6.2e-08;
Matches 36; Conservative 17; Mismatches 47; Indels 29; Gaps 6

QY 30 WC-NCPKFGGQHCEIDK-----SKTCYEGNGHFYRGKASTDTMGRPCLPWN 76
Db 225 WCYTDPQIEREFCDLPRCGSEAPQREATSVSCFRKGEGYRGANTTTAGVPCQRWDA 284
QY 77 ATVLQQTVAHRSALQLGLG---KHNYCRNPDNRRPWCY-----VOVGLKPLVQEC--- 126
Db 285 -----QIPHQRFTPEKYACKDLRENFCEPNPGSEAPWCFTLPCRMVGFYQIRCTDD 339
QY 127 -MWHDCADG 134
Db 340 VREQDCYHG 348

RESULT 41
Q9AV69
ID Q9AV69 PRELIMINARY; PRT; 930 AA.

```

```

AC Q8AV69;
DT 01-VAR-2003 (T-EMBLrel. 23, Created)
DT 01-VAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Receptor tyrosine kinase Xror2.
GN XRO2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
CX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22286220; PubMed=12399314;
RA Hikasa H., Shibata M., Hiratani I., Taira M.;
RT "The Xenopus receptor tyrosine kinase Xror2 modulates morphogenetic
RT movements of the axial mesoderm and neuroectoderm via Wnt
RT signalling.";
RL Development 129:5227-5239 (2002).
DR EMBL; AB087137; BAC16209.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0004888; F:transmembrane receptor activity; IEA.
DR GO; GO:0007275; P:development; IEA.
DR GO; GO:0008468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000024; Fz domain.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000001; Kingle.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00051; kingle; 1.
DR Pfam; PF00069; pk_nase; 1.
DR PRINTS; PR00018; KINGLE.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000395; Kingle; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00409; IG; 1; Kinase.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS50038; Fz; 1.
DR PROSITE; PS50835; IG LIKE; 1.
DR PROSITE; PS00021; KINGLE; 1.
DR PROSITE; PS50070; KINGLE_2; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW Kinase.
SQ SEQUENCE 930 AA; 104081 MW; C68454572411A8B6 CRC64;

Query Match
Best Local Similarity 19.0%; Score 150.5; DB 13; Length 930;
Matches 37; Conservative 18; Mismatches 44; Indels 33; Gaps 8;

QY 11 CDCLNGGTGVSNKY-----FNNHWNC-----PKFGQHC-----EIDSK 48
DB 252 CEVLNDLC-ROEYNIARSNPLILMLPNCLEELPLPESPAANCMEIGIPVEKLNRYQ 310
QY 49 TCYEGNGHYRKASTDGTGRCLPWSNATVLOQTYHAH---RSDALQGLGKXNYCNP 105
DB 311 QCYNGTGTDIRGSVTSKSGHCQCPWS-----HQVPHGHSUSNADYPEIG-GGHSYCRNP 364
QY 106 DNERR-PWCYVQ 116
DB 365 GGQMEGPWCFTQ 376

```

```

RESULT 42
Q7ZTN9 PRELIMINARY; PRT; 709 AA.
AC Q7ZTN9;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Similar to macrophage stimulating 1 (Hepatocyte growth
DE factor-like).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
CX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC044008; A344008.1; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0003809; F:thrombin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0007596; P:blood coagulation; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kingle.
DR InterPro; IPR003044; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR003966; Peptidase_S1A_pr.
DR Pfam; PF00051; kingle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kingle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KINGLE; 1; 4.
DR PROSITE; PS50070; KINGLE_2; 4.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
SQ SEQUENCE 709 AA; 81123 MW; 51C44D8EBD03B76A CRC64;

Query Match
Best Local Similarity 18.8%; Score 149; DB 13; Length 709;
Matches 38; Conservative 14; Mismatches 42; Indels 30; Gaps 6;

QY 30 WC-----NCPKFGQHCIDK-----CHITKIEKQIRISNITSTCFKGEYRGKANTTSGIPQ 72
DB 250 WCVTDPNVERP-----CHITKIEKQIRISNITSTCFKGEYRGKANTTSGIPQ 305
QY 73 PWSNATVLOQTYHAHSDALQ---GLGKXNYCRPNRRPWCYVQGLKPLVQECMVH 129
DB 306 RWDT-----QAPHVHRFLPEKYPCKGLDE-NYCRNPVGSEAPWCFTTLKMRMAYCFQIK 359
QY 130 DCAD 133
DB 360 RCTD 363

RESULT 43
Q8KQ08 PRELIMINARY; PRT; 801 AA.
AC Q8KQ08;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)

```

```

01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAUSBERG R.;
RA Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RL -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC EMBL; BC030848; AAH30848.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0004889; F:transmembrane receptor activity; IEA.
DR GO; GO:0007275; F:development; IEA.
DR GO; GO:0006468; F:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000024; Fz domain.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR InterPro; IPR008266; Tyr_pkinase_AS.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS0038; FZ; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW Hypothetical protein, ATP-binding, Glycoprotein; Kringle; Transferase.
FT NON TER 1
SQ SEQUENCE 801 AA; 89201 MW; 3A5928326C8B885D CRC64;

Query Match 18.8%; Score 149; DB 11; Length 801;
Best Local Similarity 36.5%; Pred. No. 1.3e-07;
Matches 42; Conservative 11; Mismatches 36; Indels 26; Gaps 9;

QY 32 NC-----PKFGGQHCEIDKSKTCYEGNGHFGKASTDTMGRCLPWN SATVLQQT YH 85
DB 156 NCMRIGIPAERLGRYH-----QCYNAGADYRGMASTTKSGHQCPW-----ALQHP-H 203

QY 86 AHR-----SDALQGLGKHNYCRNPDRNR--PWCYVQVGLKPLVQECMVHDC A--DG 134
DB 204 SHRLSSTPEPELG-GGHAYCRNPGQMGWPCFTQ-NKNVRVELCDVPPCSPRDG 256

RESULT 44
Q8C3W2 ID Q8C3W2 PRELIMINARY; PRT; 944 AA.
AC Q8C3W2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Receptor tyrosine kinase-like orphan receptor 2.
GN ROR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Heart;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RT "Analysis of the mouse transcriptome based on functional annotation of
  60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK084752; BAC39273.1; -.
DR MGD; MGI:11347521; Ror2.
DR GO; GO:0001501; P:skeletal development; IMP.
DR InterPro; IPR000024; Fz domain.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR InterPro; IPR008266; Tyr_pkinase_AS.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000395; Kringle; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGc2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS0038; FZ; 1.
DR PROSITE; PS0035; IG LIKE; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
SQ SEQUENCE 944 AA; 105037 MW; 8708ADD4CB1B1F36 CRC64;

Query Match 18.8%; Score 149; DB 11; Length 944;
Best Local Similarity 36.5%; Pred. No. 1.6e-07;
Matches 42; Conservative 11; Mismatches 36; Indels 26; Gaps 9

QY 32 NC-----PKFGGQHCEIDKSKTCYEGNGHFGKASTDTMGRCLPWN SATVLQQT YH 85
DB 299 NCMRIGIPAERLGRYH-----QCYNAGADYRGMASTTKSGHQCPW-----ALQHP-H 346

QY 86 AHR-----SDALQGLGKHNYCRNPDRNR--PWCYVQVGLKPLVQECMVHDC A--DG 134
DB 347 SHRLSSTPEPELG-GGHAYCRNPGQMGWPCFTQ-NKNVRVELCDVPPCSPRDG 399

RESULT 45
Q8BSP6 ID Q8BSP6 PRELIMINARY; PRT; 944 AA.
AC Q8BSP6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Receptor tyrosine kinase-like orphan receptor 2.
GN ROR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Forelimb;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;

```

"Analysis of the mouse transcriptome based on functional annotation of

```

RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK031112; BAC27258.1; -.
DR MGD; MGI:1347521; Ror2.
DR GO; GO:0001501; P:skeletal development; IMP.
DR InterPro; IPR000024; Pz domain.
DR InterPro; IPR000359; Ig-like.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000395; Kringle; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00220; S_TK; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00338; FZ; 1.
DR PROSITE; PS00835; IG LIKE; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
SQ SEQUENCE 944 AA; 105053 MW; 1BBD416BE3170401 CRC64;

Query Match 18.8%; Score 149; DB 11; Length 944;
Best Local Similarity 36.5%; Pred. No. 1.6e-07;
Matches 42; Conservative 11; Mismatches 36; Indels 26; Gaps 9;

QY 32 NC-----PKFVGQHCIDKSKTCYEGNGHYRGKASTDTMGRCPLPWN SATVLQQTYY 85
DB 299 NCWIGIPAEIRLGRYH-----QCYNGSGADYRGMASTTKSGHCQCPW----ALQHP-H 346

QY 86 AHR---SDALQLGLGKHNYCRPNRRR--PMCYVQVGLKPLVQECMVHDC--DG 134
DB 347 SRLSLSTEPELG-GHAYCRNPGCGMEGFCFTO--NKNVRVELCDVPPCSPRDG 399

RESULT 46
Q9UIR8
ID Q9UIR8 PRELIMINARY; PRT; 105 AA.
AC Q9UIR8
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Apolipoprotein(a) (Fragment).
GN APOA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=21181705; PubMed=11285247;
RX Ogorelkova M., Kraft H.G., Enholm C., Utermann G.;
RT "Single nucleotide polymorphisms in exons of the apo(a) kringles IV
RT types 6 to 10 domain affect Lp(a) plasma concentrations and have
RT different patterns in Africans and Caucasians.";
RL Hum. Mol. Genet. 10:815-824(2001).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF158661; AAF03679.1; -.
DR EMBL; AF158660; AAF03679.1; JOINED.
DR HSSP; P00747; 2PK4.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR ProDom; PD000001; (Fragment).
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kringle; Lipoprotein.
FT NON_TER 1
FT NON_TER 113
SQ SEQUENCE 113 AA; 12697 MW; 51D4461D9C66312E CRC64;

Query Match 18.5%; Score 147; DB 4; Length 113;
Best Local Similarity 36.0%; Pred. No. 2.6e-08;
Matches 31; Conservative 14; Mismatches 33; Indels 8; Gaps 4

QY 50 CYEGNGHYRGKASTDTMGRCPLPWN SATVLQQTYYAHRS DALQLGLGKHNYCRPNRR 109
DB 11 CYHGDGRSYRGISSTVTGRTCSQSWSS--MIPHWQRTPEYYPNGGLTENYCRPDSGK 67

QY 110 RWCYVQVGLKPLV--QECMVHDCAD 133
DB 68 SPWCYT---MDPNVRWEYCNLTQC 88

RESULT 47
Q9UIR6
ID Q9UIR6 PRELIMINARY; PRT; 113 AA.
AC Q9UIR6
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Apolipoprotein(a) (Fragment).
GN APOA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=21181705; PubMed=11285247;
RX Ogorelkova M., Kraft H.G., Enholm C., Utermann G.;
RT "Single nucleotide polymorphisms in exons of the apo(a) kringles IV
RT types 6 to 10 domain affect Lp(a) plasma concentrations and have
RT different patterns in Africans and Caucasians.";
RL Hum. Mol. Genet. 10:815-824(2001).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF158661; AAF03679.1; -.
DR EMBL; AF158660; AAF03679.1; JOINED.
DR HSSP; P00747; 2PK4.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR ProDom; PD000001; (Fragment).
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kringle; Lipoprotein.
FT NON_TER 1
FT NON_TER 113
SQ SEQUENCE 113 AA; 12697 MW; 51D4461D9C66312E CRC64;

Query Match 18.7%; Score 148; DB 4; Length 105;
Best Local Similarity 38.1%; Pred. No. 1.8e-08;
Matches 32; Conservative 8; Mismatches 36; Indels 8; Gaps 3

QY 50 CYEGNGHYRGKASTDTMGRCPLPWN SATVLQQTYYAHRS DALQLGLGKHNYCRPNRR 109
DB 11 CYHGDGRSYRGISSTVTGRTCSQSWSSMT--PHMWQRTPEYYPNGGLTENYCRPDAEI 67

QY 110 RWCYVQVGLKPLV--QECMVHDC 131
DB 68 SPWCYT---MDPNVRWEYCNLTQC 88

RESULT 48
Q9UIR6
ID Q9UIR6 PRELIMINARY; PRT; 113 AA.
AC Q9UIR6
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Apolipoprotein(a) (Fragment).
GN APOA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=21181705; PubMed=11285247;
RX Ogorelkova M., Kraft H.G., Enholm C., Utermann G.;
RT "Single nucleotide polymorphisms in exons of the apo(a) kringles IV
RT types 6 to 10 domain affect Lp(a) plasma concentrations and have
RT different patterns in Africans and Caucasians.";
RL Hum. Mol. Genet. 10:815-824(2001).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF158661; AAF03679.1; -.
DR EMBL; AF158660; AAF03679.1; JOINED.
DR HSSP; P00747; 2PK4.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR ProDom; PD000001; (Fragment).
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Glycoprotein; Kringle; Lipoprotein.
FT NON_TER 1
FT NON_TER 113
SQ SEQUENCE 113 AA; 12697 MW; 51D4461D9C66312E CRC64;

Query Match 18.5%; Score 147; DB 4; Length 113;
Best Local Similarity 36.0%; Pred. No. 2.6e-08;
Matches 31; Conservative 14; Mismatches 33; Indels 8; Gaps 4

QY 50 CYEGNGHYRGKASTDTMGRCPLPWN SATVLQQTYYAHRS DALQLGLGKHNYCRPNRR 109
DB 11 CYHGDGRSYRGISSTVTGRTCSQSWSS--MIPHWQRTPEYYPNGGLTENYCRPDSGK 67

QY 110 RWCYVQVGLKPLV--QECMVHDCAD 133
DB 68 SPWCYT---MDPNVRWEYCNLTQC 88

```

DB 58 QPWCYT---TDPCTVWEYCNLTQCS 90

RESULT 48

Q90865 PRELIMINARY; PRT; 704 AA.

AC Q90865; 704 AA.

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Hepatocyte growth factor-like/macrophage stimulating protein.

GN HGF1/MSP.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OX NCBI\_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RX MEDLINE=96029010; PubMed=7554499;

RA Thery C., Sharpe M.J., Batley S.J., Stern C.D., Gherardi E.;

RT "Expression of HGF/SF, HGF1/MSP and c-met suggests new functions

RT during early chick development.";

RL Dev. Genet. 17:90-101(1995).

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.

DR EMBL; X84043; CA58862.1; -.

DR HSSP; P00747; 1CEA.

DR MEROPS; S01.977; -.

DR GO; GO:0004263; F:chymotrypsin activity; IEA.

DR GO; GO:0008233; F:peptidase activity; IEA.

DR GO; GO:0004295; F:trypsin activity; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR009003; Cys Ser trypsin.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR003014; PAN.

DR InterPro; IPR003609; Pan app.

DR InterPro; IPR001254; Peptidase\_S1.

DR InterPro; IPR001314; Peptidase\_S1A.

DR Pfam; PF00051; kringle; 4.

DR Pfam; PF00024; PAN; 1.

DR PRINTS; PR00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR ProDom; PD000395; Kringle; 4.

DR SMART; SMART; SMC0473; PAN AP; 1.

DR SMART; SMART; SMC0020; Tryp SPC; 1.

DR PROSITE; PS00021; KRINGLE\_1; 4.

DR PROSITE; PS00070; KRINGLE\_2; 4.

DR PROSITE; PS00240; TRYPsin\_DOM; 1.

KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.

SQ SEQUENCE 704 AA; 79341 MW; CAB0D8CC41367C37 CRC64;

Query Match 18.5%; Score 146.5; DB 13; Length 704;

Best Local Similarity 36.9%; Pred. No. 2.1e-07;

Matches 41; Conservative 10; Mismatches 39; Indels 21; Gaps 7;

QY 27 NIHWNCPPKFGQHCIEDKSTCVGNGHFYRGKASTDTMGRPCLPNSATVLQQTVHA 86

DB 93 NIHYDLYQKK-----DPLRECIYANGTSYGRDTRTERGLRQCHQWATP-----HD 139

QY 87 HR-SDALQGLGKHNYCRNPD-NRRPWCYVQVGLKPLV-QECMVHDCAD 133

DB 140 HRFLPSLRNGL-EENYCRNPDNRKRGPCWYCT---VDPNVRHOSGCIKKCED 186

Query Match 18.5%; Score 146.5; DB 13; Length 704;

Best Local Similarity 36.9%; Pred. No. 2.1e-07;

Matches 41; Conservative 10; Mismatches 39; Indels 21; Gaps 7;

QY 27 NIHWNCPPKFGQHCIEDKSTCVGNGHFYRGKASTDTMGRPCLPNSATVLQQTVHA 86

DB 93 NIHYDLYQKK-----DPLRECIYANGTSYGRDTRTERGLRQCHQWATP-----HD 139

QY 87 HR-SDALQGLGKHNYCRNPD-NRRPWCYVQVGLKPLV-QECMVHDCAD 133

DB 140 HRFLPSLRNGL-EENYCRNPDNRKRGPCWYCT---VDPNVRHOSGCIKKCED 186

RESULT 49

Q90ZNG PRELIMINARY; PRT; 709 AA.

AC Q90ZNG; 709 AA.

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Hepatocyte growth factor-like 1.

GN MST1 OR HGFL1.

OS Brachydanio rerio (Zebrafish) (Danio rerio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;

OC Cyprinidae; Danio.

OX NCBI\_TaxID=7955;

RN [1]

RP SEQUENCE FROM N.A.

RA Bassett D.I., Wilson S.W.;

RT "Early expression of zebrafish Hepatocyte Growth Factor-Like 1

RT suggests a conserved role in vertebrate neural induction.";

RL Submitted (APR-2001) to the EMBL/GenBank/DBSJ databases.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.

DR EMBL; AF370035; AAK54207.1; -.

DR HSSP; P00761; 1ANI.

DR ZFIN; ZDB-GENE-020806-3; mst1.

DR GO; GO:0004263; F:chymotrypsin activity; IEA.

DR GO; GO:0008233; F:peptidase activity; IEA.

DR GO; GO:0004295; F:trypsin activity; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR009003; Cys Ser trypsin.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR003014; PAN.

DR InterPro; IPR003609; Pan app.

DR InterPro; IPR001254; Peptidase\_S1.

DR InterPro; IPR001314; Peptidase\_S1A.

DR Pfam; PF00051; kringle; 4.

DR Pfam; PF00024; PAN; 1.

DR PRINTS; PR00089; trypsin; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR ProDom; PD000395; Kringle; 4.

DR SMART; SMART; SMC0473; PAN AP; 1.

DR SMART; SMART; SMC0020; Tryp SPC; 1.

DR PROSITE; PS00021; KRINGLE\_1; 4.

DR PROSITE; PS00070; KRINGLE\_2; 4.

DR PROSITE; PS00240; TRYPsin\_DOM; 1.

KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.

SQ SEQUENCE 709 AA; 81271 MW; 990736C5DB73A20 CRC64;

Query Match 18.5%; Score 146.5; DB 13; Length 709;

Best Local Similarity 35.9%; Pred. No. 2.1e-07;

Matches 37; Conservative 12; Mismatches 33; Indels 21; Gaps 6

QY 41 HCEIDKSKT---CYEGNGHFYRGKASTDTMGRPCLPNSATVLQQTVHA 91

DB 98 NCDLYEMKVVYRKCIYKGGEDYRGKVTTSIGTCQQQWS-----KFPDHRWTSPATNG 152

QY 92 LQGLGKHNYCRNPD-NRRPWCYVQVGLKPLVQECMVHDCAD 133

DB 153 LEL-----NYCRNPDGRIGPCWYCT-TDPERRYESCNIPOCKD 189

RESULT 50

Q17576 PRELIMINARY; PRT; 902 AA.

AC Q17576; 902 AA.

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE KIN-8 protein (Receptor tyrosine kinase).

GN COIG6.8 OR KIN-8.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI\_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.



RA Berks M.;  
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99069613; PubMed=9851916;  
RA none;  
RT "Genome sequence of the nematode C.elegans: A platform for  
RT investigating biology."  
RL Science 282:2012-2018 (1998).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Mortimore B.J.;  
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=N2;  
RA Koga M., Take-uchi M., Tameishi T., Ohshima Y.;  
RT "Control of DAF-7 TGF expression and neuronal process development by a  
RT receptor tyrosine kinase KIN-8 in C. elegans."  
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  
DR EMBL; Z35595; CAA84639.2; -;  
DR EMBL; Z47808; CAA84639.2; JOINED.  
DR EMBL; Z47808; CAB61003.1; -;  
DR EMBL; Z35595; CAB61003.1; JOINED.  
DR EMBL; AJ132947; CAC29085.1; -;  
DR FIR; T18840; T18840.  
DR HSP; P11362; 1FGK.  
DR WormRep; C01G6.8b; C224774.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR GO; GO:0004888; F:transmembrane receptor activity; IEA.  
DR GO; GO:0007275; P:development; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR000024; Fz domain.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003598; Ig\_c2.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR001245; Tyr\_kinase.  
DR InterPro; IPR008266; Tyr\_kinase\_AS.  
DR Pfam; PF00047; ig; 1.  
DR Pfam; PF00051; kringle; 1.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00018; KRINGLE.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000395; Kringle; 1.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR SMART; SM00408; IGC2; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00219; TyrcK; 1.  
DR PROSITE; PS00338; FZ; 1.  
DR PROSITE; PS00835; IG LIKE; 1.  
DR PROSITE; PS00021; KRINGLE 1; FALSE\_NEG.  
DR PROSITE; PS00070; KRINGLE\_2; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
KW ATP-binding; Immunoglobulin domain; Kinase; Kringle; Transferase;  
KW Tyrosine-protein kinase.  
SQ SEQUENCE 902 AA; 101038 MW; 2A03D76D07C552B5 CRC64;

Query Match 18.2%; Score 144.5; DB 5; Length 902;  
Best Local Similarity 25.5%; Pred. NO. 4.6e-07;  
Matches 42; Conservative 25; Mismatches 55; Indels 43; Gaps 9;  
QY 1 SNELHQVPS---NCDCLNGGTCVSNKYFSNIH-----WCNCPKKFGQHC----42  
DB 248 SDNNNQIVSICKHDCDVIQNDCEPSLALAAQHVLGDTPKALFPLCSRLSSTSNCPVM 307  
QY 43 -----EIDK---SKTCYGNHGHYRGKASTDTMGRPCLPWN SATVLTQOTVHAHRS 89

Search completed: May 25, 2004, 14:57:11  
Job time : 23.362 secs

Db 308 STALOSSPVAEVNRGHLTHWCYVNSGTQYEGTVAGTSSGKQCAPWIDT--SRDENVRFP 365  
QY 90 DALQLGLGKHNYCRNPDRR-RPWCVYQVGLKPLVQE--CMVHDC 131  
DB 366 PEL---MNSKNYCRNPGGKSRPCY-----SKPMGQEBYCDVPQC 403

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2004, 14:43:00 ; Search time 33.0105 Seconds

(without alignments)  
1155.508 Million cell updates/sec

Title: US-09-880-503-4

Perfect score: 793

Sequence: 1 SNELHQVPSNCCLNGGTCV.....QVGLKPLVQEWVHDCADGK 135

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : A\_Geneseq\_29Jan04.\*

1: Geneseq1980s.\*

2: Geneseq1990s.\*

3: Geneseq2000s.\*

4: Geneseq2001s.\*

5: Geneseq2002s.\*

6: Geneseq2003as.\*

7: Geneseq2003bs.\*

8: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	793	100.0	135	5	AAE16545 Human uro
2	793	100.0	143	5	AAE16549 Human uro
3	793	100.0	337	4	AAE75492 Human col
4	793	100.0	337	5	ABP41795 Human ova
5	793	100.0	403	5	AAE16547
6	793	100.0	411	1	AAE50871
7	793	100.0	411	2	AAE06244 Urokinase
8	793	100.0	411	2	AAE05117 UK-S3 as
9	793	100.0	411	2	AAE13634 Human nat
10	793	100.0	411	2	AAE10057
11	793	100.0	411	2	AAE10058
12	793	100.0	411	2	AAE40325
13	793	100.0	411	2	AAE62997
14	793	100.0	411	2	AAE63008
15	793	100.0	411	2	AAE62999
16	793	100.0	411	2	AAE62994
17	793	100.0	411	2	AAE63003
18	793	100.0	411	2	AAE63001
19	793	100.0	411	2	AAE63006
20	793	100.0	411	2	AAE62992
21	793	100.0	411	2	AAE63002
22	793	100.0	411	2	AAE63007
23	793	100.0	411	2	AAE62995
24	793	100.0	411	2	AAE63000
25	793	100.0	411	2	AAE63004

26	793	100.0	411	2	AAE62996
27	793	100.0	411	2	AAE63005
28	793	100.0	411	2	AAE62991
29	793	100.0	411	2	AAE62998
30	793	100.0	411	2	AAE62993
31	793	100.0	411	2	AAE92926
32	793	100.0	411	3	AAE92836 Urokinase
33	793	100.0	411	5	AAE16544 Human uro
34	793	100.0	412	2	AAE10334
35	793	100.0	424	2	AAW24579
36	793	100.0	430	2	AAW24578
37	793	100.0	431	1	AAE50114 Sequence
38	793	100.0	431	1	AAE60783 Human uro
39	793	100.0	431	1	AAE70258 Sequence
40	793	100.0	431	1	AAE71491 Modified
41	793	100.0	431	1	AAE71663 Modified
42	793	100.0	431	1	AAE71698 Modified
43	793	100.0	431	1	AAE71699 Modified
44	793	100.0	431	1	AAE80430 Deduced A
45	793	100.0	431	1	AAE81204 Pro-uroki
46	793	100.0	431	1	AAE91886 Sequence
47	793	100.0	431	1	AAE92119 Natural h
48	793	100.0	431	1	AAE94764 Non-glyco
49	793	100.0	431	2	AAE07112 Human pro
50	793	100.0	431	2	AAE04253 Human pro
51	793	100.0	431	2	AAE63141 Full leng
52	793	100.0	431	2	AAE47903 Pro-uroki
53	793	100.0	431	2	AAE33199 Human uri
54	793	100.0	431	2	AAE33198 Human uri
55	793	100.0	431	2	AAE33121 Human uri
56	793	100.0	431	3	AAE50869 Human uro
57	793	100.0	431	3	AAE99591 Human pla
58	793	100.0	431	4	AAE84605 Amino aci
59	793	100.0	431	5	AAE17128 Human uro
60	793	100.0	431	5	AAE99228 Human pla
61	793	100.0	431	5	AAE99231 Human pla
62	793	100.0	431	5	AAE99229 Human pla
63	793	100.0	431	5	AAE99230 Human pla
64	793	100.0	431	5	AAE79460 sc-uPA. 1
65	793	100.0	431	6	AAE37128 Human uro
66	793	100.0	431	6	ABE55855 Human uro
67	793	100.0	431	6	ABE56547 Lung canc
68	793	100.0	431	6	ABE56708 Lung canc
69	793	100.0	431	6	ABU11076 Human uro
70	793	100.0	431	6	ABR2137 Human cer
71	793	100.0	431	7	ADD46429 Human pro
72	793	100.0	434	2	AAE20537 Amidated
73	793	100.0	434	2	AAE20538 Amidated
74	793	100.0	436	2	AAE20536 Amidated
75	791	99.7	431	7	AAE25745 Human pro

#### ALIGNMENTS

RESULT 1

AAE16545 AAE16545 standard; protein; 135 AA.

XX AAE16545;

AC AAE16545;

XX 09-APR-2002 (first entry)

DT Human urokinase-type plasminogen activator amino terminal fragment (ATF).

XX Human; urokinase-type plasminogen activator; uPA; therapy; hypertension; stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder; microvascular occlusion; angiogenic disorder; pulmonary fibrosis; aschma; tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing; clotting disorder; uterine contraction disorder; respiratory disease; adult respiratory distress syndrome; amino terminal fragment; ATF; male impotence.



XX OS Homo sapiens.  
 XX FN WO200122920-A2.  
 XX PD 05-APR-2001.  
 XX PF 28-SEP-2000; 2000WO-US026524.  
 XX PR 29-SEP-1999; 99US-0157137P.  
 XX PR 03-NOV-1999; 99US-0163280P.  
 XX XX  
 XX PA (HUMA-) HUMAN GENOME SCI INC.  
 XX PI Ruben SM, Barash SC, Birse CE, Rosen CA;  
 XX DR WPI; 2001-235357/24.  
 XX DR N-PSDB; AAH34897.  
 XX PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
 XX PT useful for preventing, diagnosing and/or treating colorectal cancers.  
 XX PS Claim 11; Page 7707-7708; 9803pp; English.  
 XX XX  
 CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where the  
 CC proteins are collectively known as colon cancer antigens. The colon  
 CC cancer antigens have cytostatic activity and can be used in gene therapy  
 CC and vaccine production. N and P may be used in the prevention, diagnosis  
 CC and treatment of diseases associated with inappropriate P expression. For  
 CC example, N and P may be used to treat disorders associated with decreased  
 CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of P by expressing inactive proteins or to  
 CC supplement the patient's own production of P. Additionally, N may be used  
 CC to produce the colon cancer-associated PS, by inserting the nucleic acids  
 CC into a host cell and culturing the cell to express the proteins. N and P  
 CC can be used in the prevention, diagnosis and treatment of colorectal  
 CC carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent  
 CC sequences used in the exemplification of the present invention. N.B.  
 CC Pages 666 to 682 and page 7053 of the sequence listing were missing at  
 CC time of publication. Meaning no sequences are present for SEQ ID NO:1027  
 CC to 1052, 7921 and 7922  
 XX XX  
 SQ Sequence 337 AA;

Query Match 100.0%; Score 793; DB 4; Length 337;  
 Best Local Similarity 100.0%; Pred. No. 7.3e-54;  
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFFYRG 60  
 DB 27 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFFYRG 86  
 QY 61 KASDTMTGRPCLPNSATVLOQTYHAHRS DALQLGLGHNYCRNPDRRPPWCYVOYGLK 120  
 DB 87 KASDTMTGRPCLPNSATVLOQTYHAHRS DALQLGLGHNYCRNPDRRPPWCYVOYGLK 146  
 QY 121 PLVQECMVHDCADGK 135  
 DB 147 PLVQECMVHDCADGK 161

RESULT 4  
 ABP41795  
 ID ABP41795 standard; protein; 337 AA.  
 XX AC ABP41795;  
 XX XX  
 XX DT 22-AUG-2002 (first entry)  
 XX DE Human ovarian antigen HVVCE79, SEQ ID NO:2927.  
 XX KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;

KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
 KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;  
 KW inflammatory condition; immune disorder; blood disorder;  
 KW cardiovascular disorder; respiratory disorder; neurological disorder;  
 KW gastrointestinal disorder; urinary system disorder; drug screening;  
 KW gene therapy; chromosome mapping; forensic analysis;  
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
 KW antiinflammatory; gynaecological; reproductive; chromosome 10q24.  
 XX OS Homo sapiens.  
 XX XX  
 XX FN WO200200677-A1.  
 XX PD 03-JAN-2002.  
 XX XX  
 XX PF 07-JUN-2001; 2001WO-US018569.  
 XX XX  
 XX PR 07-JUN-2000; 2000US-0209467P.  
 XX XX  
 XX XX (HUMA-) HUMAN GENOME SCI INC.  
 XX PI Birse CE, Rosen CA;  
 XX DR WPI; 2002-147878/19.  
 XX DR N-PSDB; ABQ54872.  
 XX PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
 XX PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian  
 XX PT cancer), immune disorders, cardiovascular disorders and neurological  
 XX PT diseases.  
 XX PS Claim 11; SEQ ID NO 2927; 2922pp; English.  
 XX XX  
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-  
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
 CC to the sequences of the invention. The invention additionally relates to  
 CC recombinant vectors and host cells comprising human ovarian antigen  
 CC polynucleotides, antibodies against human ovarian antigens, and the use  
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
 CC treating, prognosing or preventing various ovary and/or breast-related  
 CC disorders. Such conditions include ovarian cancer and breast cancer, and  
 CC metastatic tumours of ovarian or breast origin, reproductive system  
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine  
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
 CC vaginitis), immune disorders (e.g., congenital and acquired  
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
 CC respiratory disorders, neurological disorders, gastrointestinal disorders  
 CC and urinary system disorders. Ovarian antigen polypeptides and  
 CC polynucleotides may also be used in screening for compounds which  
 CC modulate ovarian antigen expression or activity. The polynucleotides may  
 CC further be used for gene therapy, chromosome mapping, in the  
 CC identification of individuals and in forensic analysis, and the  
 CC polypeptides may be used as food additives or to prepare antibodies  
 CC useful in disease diagnosis, drug targeting and phenotyping. The present  
 CC sequence represents a human ovarian antigen of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX XX  
 SQ Sequence 337 AA;

Query Match 100.0%; Score 793; DB 5; Length 337;  
 Best Local Similarity 100.0%; Pred. No. 7.3e-54;  
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0  
 QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFFYRG 60  
 DB 27 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFFYRG 86

QY 61 KASTDTMGRPCLPWNATVLOQTYHAHRSALQLGLGKHNYCRNPDNRPRPWCYVQVGLK 120  
DB 87 KASTDTMGRPCLPWNATVLOQTYHAHRSALQLGLGKHNYCRNPDNRPRPWCYVQVGLK 146  
QY 121 PLVQECMVHDCADGK 135  
DB 147 PLVQECMVHDCADGK 161  
RESULT 5  
AAE1547  
ID AAE16547 standard; protein; 403 AA.  
XX AAE16547;  
XX  
DT 09-APR-2002 (first entry)  
DE Human urokinase-type plasminogen activator scuPA delta136-143 mutant.  
XX  
XX Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;  
KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;  
KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;  
KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;  
KW clotting disorder; uterine contraction disorder; respiratory disease;  
KW male impotence; adult respiratory distress syndrome; scuPA delta136-143;  
KW single chain urokinase; mutant; mutein.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN WO200197752-A2.  
XX  
PD 27-DEC-2001.  
XX  
PF 13-JUN-2001; 2001WO-US018976.  
XX  
PR 20-JUN-2000; 2000US-0212874P.  
XX  
XX (UYPE-) UNIV PENNSYLVANIA.  
PA Cines DB, Higazi AA;  
XX WPI; 2002-122240/16.  
DR N-PSDB; AAD27080.  
XX  
XX Composition for modulating muscle cell and tissue contractility for  
PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,  
PT comprising domains from urokinase-type plasminogen activator.  
XX  
PS Claim 22; Fig 1F; 117pp; English.  
XX  
XX The invention relates to a composition comprising one or more domains of  
CC urokinase-type plasminogen activator (uPA). The composition is used to  
CC modulate the contractility and angiogenic activity of a mammalian muscle,  
CC endothelial cell or tissue. The composition is used for treating stroke,  
CC hypotension, hypertension, atherosclerosis, heart attack, microvascular  
CC occlusions, thrombotic microangiopathies, surgically induced thrombotic  
CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell  
CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,  
CC diabetic retinopathy, wound healing, clotting disorder, uterine  
CC contraction disorder, male impotence, respiratory disease or condition  
CC such as asthma, adult respiratory distress syndrome, primary pulmonary  
CC hypertension, microvascular thrombotic occlusion, and a disorder  
CC associated with chronic intrapulmonary fibrin formation. The present  
CC sequence is human urokinase-type plasminogen activator (uPA) single chain  
CC urokinase (scuPA) deletion mutant designated as scuPA delta136-143  
XX  
SQ Sequence 403 AA;  
Query Match 100.0%; Score 793; DB 5; Length 403;  
Best Local Similarity 100.0%; Pred. No. 8.6e-54;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHGVSNCDLNGGTCVSNKYFNSNTHWCNCPKKGQHCEDKSKTCYEGNGHYRG 60  
DB 1 SNELHGVSNCDLNGGTCVSNKYFNSNTHWCNCPKKGQHCEDKSKTCYEGNGHYRG 60  
QY 61 KASTDTMGRPCLPWNATVLOQTYHAHRSALQLGLGKHNYCRNPDNRPRPWCYVQVGLK 120  
DB 61 KASTDTMGRPCLPWNATVLOQTYHAHRSALQLGLGKHNYCRNPDNRPRPWCYVQVGLK 120  
QY 121 PLVQECMVHDCADGK 135  
DB 121 PLVQECMVHDCADGK 135  
RESULT 6  
AAP50871  
ID AAP50871 standard; protein; 411 AA.  
XX AAP50871;  
XX  
DT 30-NOV-1991 (first entry)  
DE Sequence encoded by cDNA sequence for human urokinase zymogen (Japanese  
DE Patent Application No.37119/84).  
XX  
KW Thrombolytic agent; plasminogen activator activity; fibrin affinity;  
KW enzyme.  
XX  
OS Homo sapiens.  
PH Key Location/Qualifiers  
FT Disulfide-bond 50..131  
FT Disulfide-bond 71..113  
FT Disulfide-bond 102..126  
FT Disulfide-bond 148..279  
FT Cleavage-site 158..159  
FT /note= "potential cleavage site which generates the two-  
FT chain form from the zymogen"  
FT Disulfide-bond 189..205  
FT Disulfide-bond 197..268  
FT Disulfide-bond 293..362  
FT Disulfide-bond 325..341  
FT Disulfide-bond 352..380  
XX  
PN EP139447-A.  
XX  
PD 02-MAY-1985.  
XX  
PF 07-SEP-1984; 84EP-00306117.  
XX  
PR 13-SEP-1983; 83JP-00170354.  
PR 17-OCT-1983; 83JP-00195051.  
XX  
XX (GREG ) GREEN CROSS CORP.  
XX  
XX Kasai S, Arimura H, Mori K, Suyama T;  
DR WPI; 1985-106530/18.  
XX  
XX New urokinase zymogen - useful as thrombolytic agent.  
XX  
PS Disclosure; Page 12; 30pp; English.  
XX  
XX Zymogen AAP50871 is the inactive precursor form of human urokinase.  
CC Urokinase zymogen is cleaved into the two-chain form composed of  
CC characteristic urokinase H (molecular wt. of 30,000) and L (molecular  
CC wt. of 20,000) chains when treated with catalytic amounts of plasmin. The  
CC patentors claim a new urokinase zymogen which has mol. wt. ca. 50,000, a  
CC single chain molecular structure, and selective affinity for fibrin. It  
CC is a thrombolytic agent which manifests its plasminogen activator  
CC activity on cleavage by proteolytic enzymes (e.g. plasmin) and has higher  
CC affinity for fibrin than known forms of urokinase  
XX

```
SQ Sequence 411 AA;
Query Match
Best Local Similarity 100.0%; Score 793; DB 1; Length 411;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHN YCRNPNRRRPMWCVVQVGLK 120
DB 61 KASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHN YCRNPNRRRPMWCVVQVGLK 120
QY 121 PLVQECMVHDCADGK 135
DB 121 PLVQECMVHDCADGK 135

RESULT 7
AAR06244
ID AAR06244 standard; protein; 411 AA.
AC AC
AAAR06244;
DT 07-DEC-1990 (first entry)
DE Urokinase precursor protein.
XX Urokinase precursor; fibrinolysis; thrombolytic; cerebral thrombosis;
KW myocardial infarction.
XX Homo sapiens.
OS Homo sapiens.
FN EP390334-A.
XX 01-AUG-1990.
XX 25-JAN-1990; 90EP-00300772.
PR 27-JAN-1989; 89JP-00016406.
XX 17-MAY-1989; 89JP-00121405.
PA (GREC ) GREEN CROSS CORP.
XX Matsuda H, Ueda Y, Tamanouchi K;
PI WPI; 1990-233117/31.
DR Urokinase precursor-lipid composite - used as thrombolytic agent, having
XX prolonged half-life in the blood, enhanced bioavailability and improved
PT activity.
PS Claim 3; Fig 1; 11pp; English.
XX By forming a precursor-lipid composite, the half-life of this
CC thrombolytic agent in the blood may be increased, exhibiting improved
CC activity without abnormal acceleration of fibrinolytic activity. Compound
CC is useful as a thrombolytic agent in treatment of cerebral thrombosis,
CC myocardial infarction etc
XX Sequence 411 AA;
Query Match
Best Local Similarity 100.0%; Score 793; DB 2; Length 411;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHN YCRNPNRRRPMWCVVQVGLK 120
DB 61 KASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHN YCRNPNRRRPMWCVVQVGLK 120
QY 121 PLVQECMVHDCADGK 135
DB 121 PLVQECMVHDCADGK 135

us-09-880-503-4.rag
Query Match
Best Local Similarity 100.0%; Score 793; DB 1; Length 411;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 KASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHN YCRNPNRRRPMWCVVQVGLK 120
QY 121 PLVQECMVHDCADGK 135
DB 121 PLVQECMVHDCADGK 135

RESULT 8
AAR05117
ID AAR05117 standard; protein; 411 AA.
AC AC
AAAR05117;
DT 25-MAR-2003 (revised)
DT 04-OCT-1990 (first entry)
XX UK-S3 as encoded by PUKS3.
XX Urokinase; glycosylation.
XX Homo sapiens.
XX Key
FT Misc-difference 153 Location/Qualifiers
FT /label= synthetic mutation
FT /note= "old seq (Ileu)"
FT Misc-difference 155
FT /label= synthetic mutation
FT /note= "old seq (Pro)"
XX EP370205-A.
XX 30-MAY-1990.
XX 28-SEP-1989; 89EP-00117981.
XX 29-SEP-1988; 88JP-00245705.
XX (KYOW ) KYOWA HAKKO KOGYO KK.
XX Sasaki K, Nishi T, Yasumura S, Sato M, Itoh S;
XX WPI; 1990-165029/22.
DR N-PSDB; AAQ04486.
XX Polypeptide(s) with added carbohydrate chains - formed by modification of
PT aminoacid sequence, used to improve physio:chemical properties and/or
PT activities.
XX Disclosure; Page ?; 30pp; English.
XX The polypeptide is a deriv. of mature urokinase, designated UK-S3 which
CC has 2 amino acid substns. which result in an N-linked glycosylation site
CC giving the new protein improved stability and activity. See also AAR05113
CC -17. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-
CC 2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)
XX Sequence 411 AA;
Query Match
Best Local Similarity 100.0%; Score 793; DB 2; Length 411;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHN YCRNPNRRRPMWCVVQVGLK 120
DB 61 KASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHN YCRNPNRRRPMWCVVQVGLK 120
QY 121 PLVQECMVHDCADGK 135
DB 121 PLVQECMVHDCADGK 135
```

```

Db      121 PLVQECMVHDCADGK 135
RESULT 9
AAWI13634
ID      AAWI13634 standard; protein; 411 AA.
XX
XX
AC      AAWI13634;
XX
DT      04-JUN-1997 (first entry)
XX
DE      Human native prourokinase.
XX
KW      Human; prourokinase; hPUK; variant; half-life; increase; EGF;
KW      epidermal growth factor domain; deletion; thrombolysis; fibrinolysis.
XX
OS      Homo sapiens.
XX
FH      Key
FH      Domain
FT      10..42
FT      /label= EGF domain
FT      /note= "in Claimed variants, at least part of the EGF
FT      domain is deleted (see comments)"
FT      10..19
FT      /label= first_loop
FT      20..31
FT      /label= second_loop
FT      33..42
FT      /label= third_loop
XX
FN      EP398361-A.
XX
XX
PD      22-NOV-1990.
XX
XX
PF      18-MAY-1990; 90EP-00109472.
XX
XX
PR      18-MAY-1989; 89JP-00126433.
PR      22-FEB-1990; 90JP-00042020.
XX
XX
PA      (GREC ) GREEN CROSS CORP.
XX
XX
PI      Tanabe T, Amatsuji Y, Kasai S, Hirose M, Morita M, Kawabe H;
PI      Airmura H;
XX
XX
WFI; 1990-350146/47.
DR      N-PSDB; AAT61671.
XX
XX
Human pro-urokinase variants - deficient in loop regions of epidermal
growth factor, showing long blood half-life, as fibrinolytic agent.
XX
PS      Disclosure; Fig 1; 22pp; English.
XX
CC      New variants of human prourokinase (hPUK) comprise a hPUK deficient in
CC      (i) at least part of the first loop region of the epidermal growth factor
CC      (EGF) domain; (ii) at least part of the first loop and at least part of
CC      the second loop; or (iii) at least part of the third loop. The hPUK
CC      variants show an increased blood half-life comparable to that of the
CC      whole EGF domain-deficient hPUK variant and urokinase while retaining the
CC      same properties as those of hPUK. They have potent thrombolytic activity
CC      and very little tendency to cause spontaneous bleeding. The present
CC      sequence is that of the wild-type hPUK protein, including the EGF domain
XX
XX
Sequence 411 AA;
Query Match      100.0%; Score 793; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 8.7e-54;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKKGQHCIDKSKTCYEGNGHFYRG 60
DB      1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKKGQHCIDKSKTCYEGNGHFYRG 60
QY      61 KASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRRPWCYVQVGLK 120
DB      61 KASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRRPWCYVQVGLK 120

Query Match      100.0%; Score 793; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 8.7e-54;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      121 PLVQECMVHDCADGK 135
DB      121 PLVQECMVHDCADGK 135
XX
XX
RESULT 11
AAR10058
ID      AAR10058 standard; protein; 411 AA.
XX

```

AC AAR10058;  
 XX 18-MAR-1991 (first entry)  
 XX Pro-urokinase derivative UK-S3 with Asn(153) and Thr(155).  
 XX pro-urokinase; UK-S3; plasminogen activator; myocardial infarction;  
 KW cerebral thrombosis.  
 XX OS  
 XX Homo sapiens.  
 PN EP405285-A.  
 XX 02-JAN-1991.  
 XX 18-JUN-1990; 90EP-00111471.  
 XX 19-JUN-1989; 89JP-00156302.  
 XX (KYOW ) KYOWA HAKKO KOGYO KK.  
 XX Yasamura S, Nishi T, Ito S;  
 PI WPI; 1991-008678/02.  
 DR N-PSDB; AAQ10170.  
 XX New plasminogen activator almost identical to natural pro-urokinase - is  
 PT thrombin resistant and used for prophylaxis-treatment of cerebral  
 PT thrombosis or myocardial infarction.  
 XX Dislosure; Page 9; 84pp; English.  
 XX UK-S3 is one example of a plasminogen activator which differs from  
 CC natural human pro-urokinase at positions 153 and 155. (Leu substituted by  
 CC Asn; Pro substituted by Thr, respectively). The derivative has decreased  
 CC susceptibility to thrombin compared to natural type pro-UK and higher  
 CC specific activity. See also AAQ10168 and AAQ10169  
 XX Sequence 411 AA;  
 SQ

Query Match 100.0%; Score 793; DB 2; Length 411;  
 Best Local Similarity 100.0%; Pred. No. 8.7e-54;  
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SNELHGVPSNCDCLNGGTCVSNKYPSNIHWNCNCPKFGQHCCEIDKSKTCYEGNGHFYRG 60  
 DB 1 SNELHGVPSNCDCLNGGTCVSNKYPSNIHWNCNCPKFGQHCCEIDKSKTCYEGNGHFYRG 60  
 QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120  
 DB 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120  
 QY 121 PLVQECMVHDCADGK 135  
 DB 121 PLVQECMVHDCADGK 135

RESULT 12  
 AAR40225  
 ID AAR40225 standard; protein; 411 AA.  
 XX AAR40225;  
 XX 10-FEB-1994 (first entry)  
 XX PUK.  
 XX Pre-urokinase; thrombolytic; blood; plasmid; PUK.  
 KW OS  
 XX Homo sapiens.  
 XX Key Location/Qualifiers  
 FT Region 10. .60

/note= "Ser, Asn, Pro, Gly, or Tyr in the region 10 to 60  
 in the N-terminal of the human PUK can be replaced by  
 Thr, Pro or Ala"  
 10. .49  
 /label= EGF  
 10. .19  
 /label= Loop\_1  
 20. .31  
 /label= Loop\_2  
 33. .42  
 /label= Loop\_3  
 Misc-difference 64  
 /note= "Thr encoded by AGT (sic)"  
 Misc-difference 177  
 /note= "Thr encoded by TAC (sic)"  
 JP05192142-A.  
 PN 03-AUG-1993.  
 PD 20-JAN-1992; 92JP-00030178.  
 PF 20-JAN-1992; 92JP-00030178.  
 PR (GREC ) GREEN CROSS CORP.  
 PA WPI; 1993-277461/35.  
 DR N-PSDB; AAQ48228.  
 XX Mutant human pre-urokinase - by replacing specified aminoacid(s) in N-  
 PT terminal for providing longer half-life in blood and higher thrombolytic  
 PT ability.  
 PS Claim 1; Page 14-16; 26pp; Japanese.  
 XX Human pre-urokinase can be modified by replacing: Ser, Asn, Pro, Gly, or  
 CC Tyr in the region ranging from residue 10 to 60 in the N-terminal of the  
 CC human PUK by Thr, Pro or Ala. These mutants have a longer half-life in  
 CC the blood and a higher thrombolytic ability. For examples see (AAR47956-  
 CC R47960)  
 XX Sequence 411 AA;  
 SQ

Query Match 100.0%; Score 793; DB 2; Length 411;  
 Best Local Similarity 100.0%; Pred. No. 8.7e-54;  
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SNELHGVPSNCDCLNGGTCVSNKYPSNIHWNCNCPKFGQHCCEIDKSKTCYEGNGHFYRG 60  
 DB 1 SNELHGVPSNCDCLNGGTCVSNKYPSNIHWNCNCPKFGQHCCEIDKSKTCYEGNGHFYRG 60  
 QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120  
 DB 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120  
 QY 121 PLVQECMVHDCADGK 135  
 DB 121 PLVQECMVHDCADGK 135

RESULT 13  
 AAR62997  
 ID AAR62997 standard; protein; 411 AA.  
 XX AAR62997;  
 XX 25-MAR-2003 (revised)  
 DT 21-SEP-1995 (first entry)  
 XX Pro-urokinase mutant Gly306.  
 XX Pro-urokinase; thrombolysis; fibrin clot lysis; mutant Gly306;  
 KW reduced fibrinogenolysis; non-specific plasminogen activation;  
 FT



systemic bleeding.  
Homo sapiens.

Key Location/Qualifiers  
Disulfide-bond 11..19  
Disulfide-bond 13..31  
Disulfide-bond 33..42  
Disulfide-bond 50..131  
Disulfide-bond 71..113  
Disulfide-bond 102..126  
Disulfide-bond 148..279  
Disulfide-bond 189..205  
Disulfide-bond 197..268  
Disulfide-bond 293..362  
Domain 297..313  
Disulfide-bond 325..341  
Disulfide-bond 352..380  
WO9501427-A1.  
12-JAN-1995.  
28-JUN-1994; 94WO-US007278.  
02-JUL-1993; 93US-00087163.  
(NEWE-) NEW ENGLAND DEACONESS HOSPITAL.  
Liu J, Gurewich V;  
WPI; 1995-060991/08.  
Pro-urokinase mutants - have thrombolytic activity but reduced fibrinogenolysis activity and non-specific plasminogen activation.  
Claim 13; Fig 1; 46pp; English.  
AAR62991 is the wild type pro-urokinase, from which the new mutants described in AAR62992-R63008 were derived. These mutants retain the thrombolytic activity of the wild type protein, useful for the treatment of thromboembolism, but have a reduced fibrinogenolysis activity and non-specific plasminogen activation. The mutants can therefore be used for the lysis of fibrin clots without inducing systemic bleeding, as can be the case with the wild type protein. (Updated on 25-MAR-2003 to correct PN field.)  
Sequence 411 AA;  
Query Match 100.0%; Score 793; DB 2; Length 411;  
Best Local Similarity 100.0%; Pred. No. 8.7e-54;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 SNELHVPNSCDCLNGTGVSNKYFSNIHWCNPKKFGGQHCIDKSKTCYEGNGHFYRG 60  
1 SNELHVPNSCDCLNGTGVSNKYFSNIHWCNPKKFGGQHCIDKSKTCYEGNGHFYRG 60  
61 KASTDTMGRPCLPWSATVLTQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120  
61 KASTDTMGRPCLPWSATVLTQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120  
121 PLVQECMVHDCADGK 135  
121 PLVQECMVHDCADGK 135

RESULT 14  
AAR63008  
ID AAR63008 standard; protein; 411 AA.  
XX  
AC AAR63008;  
XX

DT 25-MAR-2003 (revised)  
DT 21-SEP-1995 (first entry)  
XX  
DE Pro-urokinase mutant Ser175 His187 His300 Ala301 His313.  
XX  
KW Pro-urokinase; thrombolysis; fibrin clot lysis; reduced fibrinogenolysis;  
KW non-specific plasminogen activation; systemic bleeding;  
KW mutant Ser175 His187 His300 Ala301 His313.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Disulfide-bond 11..19  
FT Disulfide-bond 13..31  
FT Disulfide-bond 33..42  
FT Disulfide-bond 50..131  
FT Disulfide-bond 71..113  
FT Disulfide-bond 102..126  
FT Disulfide-bond 148..279  
FT Disulfide-bond 189..205  
FT Disulfide-bond 197..268  
FT Disulfide-bond 293..362  
FT Domain 297..313  
FT Disulfide-bond 325..341  
FT Disulfide-bond 352..380  
FN WO9501427-A1.  
XX  
PD 12-JAN-1995.  
XX  
PF 28-JUN-1994; 94WO-US007278.  
XX  
PR 02-JUL-1993; 93US-00087163.  
XX  
PA (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.  
XX  
PI Liu J, Gurewich V;  
XX  
DR WPI; 1995-060991/08.  
XX  
PT Pro-urokinase mutants - have thrombolytic activity but reduced fibrinogenolysis activity and non-specific plasminogen activation.  
XX  
PS Claim 13; Fig 1; 46pp; English.  
XX  
CC AAR62991 is the wild type pro-urokinase, from which the new mutants described in AAR62992-R63008 were derived. These mutants retain the thrombolytic activity of the wild type protein, useful for the treatment of thromboembolism, but have a reduced fibrinogenolysis activity and non-specific plasminogen activation. The mutants can therefore be used for the lysis of fibrin clots without inducing systemic bleeding, as can be the case with the wild type protein. (Updated on 25-MAR-2003 to correct PN field.)  
XX  
SQ Sequence 411 AA;  
Query Match 100.0%; Score 793; DB 2; Length 411;  
Best Local Similarity 100.0%; Pred. No. 8.7e-54;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 SNELHVPNSCDCLNGTGVSNKYFSNIHWCNPKKFGGQHCIDKSKTCYEGNGHFYRG 60  
1 SNELHVPNSCDCLNGTGVSNKYFSNIHWCNPKKFGGQHCIDKSKTCYEGNGHFYRG 60  
61 KASTDTMGRPCLPWSATVLTQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120  
61 KASTDTMGRPCLPWSATVLTQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120  
121 PLVQECMVHDCADGK 135  
121 PLVQECMVHDCADGK 135

[illegible]

Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQRCHEIDKSKTCYEGNGHFYRG 60  
 Db 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQRCHEIDKSKTCYEGNGHFYRG 60

QY 61 KASTDTMGRPCLPWNSATVLOOTYHAHRSALQGLGKHNYCRNPDNRPRPCVQVGLK 120  
 Db 61 KASTDTMGRPCLPWNSATVLOOTYHAHRSALQGLGKHNYCRNPDNRPRPCVQVGLK 120

QY 121 PLVQECMVHDCADGK 135  
 Db 121 PLVQECMVHDCADGK 135

RESULT 17  
 AAR63003  
 ID AAR63003 standard; protein; 411 AA.  
 XX AC AAR63003;  
 XX 25-MAR-2003 (revised)  
 DT 21-SEP-1995 (first entry)  
 XX DE Pro-urokinase mutant Ser175 His187 His313.  
 XX KW Pro-urokinase; thrombolysis; fibrin clot lysis; reduced fibrinogenolysis;  
 KW non-specific plasminogen activation; systemic bleeding;  
 KW mutant Ser175 His187 His313.  
 XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
 FT Disulfide-bond 11..19  
 FT Disulfide-bond 13..31  
 FT Disulfide-bond 33..42  
 FT Disulfide-bond 50..131  
 FT Disulfide-bond 71..113  
 FT Disulfide-bond 102..126  
 FT Disulfide-bond 148..279  
 FT Disulfide-bond 189..205  
 FT Disulfide-bond 197..268  
 FT Disulfide-bond 293..362  
 FT Domain 297..313  
 FT /note= "flexible loop"  
 FT Disulfide-bond 325..341  
 FT Disulfide-bond 352..380

XX PN WO9501427-A1.  
 XX PD 12-JAN-1995.  
 XX 28-JUN-1994; 94WO-US007278.  
 XX 02-JUL-1993; 93US-00087163.  
 XX (NEWB-) NEW ENGLAND DEACONESS HOSPITAL.  
 XX PI Liu J, Gurewich V;  
 XX WPI; 1995-060991/08.  
 XX PT Pro-urokinase mutants - have thrombolytic activity but reduced  
 PT fibrinogenolysis activity and non-specific plasminogen activation.  
 XX Claim 15; Fig 1; 46pp; English.

XX AAR62991 is the wild type pro-urokinase, from which the new mutants  
 CC described in AAR62992-R63008 were derived. These mutants retain the  
 CC thrombolytic activity of the wild type protein, useful for the treatment  
 CC of thromboembolism, but have a reduced fibrinogenolysis activity and non-  
 CC specific plasminogen activation. The mutants can therefore be used for  
 CC the lysis of fibrin clots without inducing systemic bleeding, as can be

CC the case with the wild type protein. (Updated on 25-MAR-2003 to correct  
 CC PN field.)  
 XX SQ Sequence 411 AA;  
 Query Match 100.0%; Score 793; DB 2; Length 411;  
 Best Local Similarity 100.0%; Pred. No. 8.7e-54; Indels 0; Gaps 0  
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQRCHEIDKSKTCYEGNGHFYRG 60  
 Db 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQRCHEIDKSKTCYEGNGHFYRG 60

QY 61 KASTDTMGRPCLPWNSATVLOOTYHAHRSALQGLGKHNYCRNPDNRPRPCVQVGLK 120  
 Db 61 KASTDTMGRPCLPWNSATVLOOTYHAHRSALQGLGKHNYCRNPDNRPRPCVQVGLK 120

QY 121 PLVQECMVHDCADGK 135  
 Db 121 PLVQECMVHDCADGK 135

RESULT 18  
 AAR63001  
 ID AAR63001 standard; protein; 411 AA.  
 XX AC AAR63001;  
 XX 25-MAR-2003 (revised)  
 DT 21-SEP-1995 (first entry)  
 XX DE Pro-urokinase mutant Ser175 His187 Ala313.  
 XX KW Pro-urokinase; thrombolysis; fibrin clot lysis; reduced fibrinogenolysis;  
 KW non-specific plasminogen activation; systemic bleeding;  
 KW mutant Ser175 His187 Ala313.  
 XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
 FT Disulfide-bond 11..19  
 FT Disulfide-bond 13..31  
 FT Disulfide-bond 33..42  
 FT Disulfide-bond 50..131  
 FT Disulfide-bond 71..113  
 FT Disulfide-bond 102..126  
 FT Disulfide-bond 148..279  
 FT Disulfide-bond 189..205  
 FT Disulfide-bond 197..268  
 FT Disulfide-bond 293..362  
 FT Domain 297..313  
 FT /note= "flexible loop"  
 FT Disulfide-bond 325..341  
 FT Disulfide-bond 352..380

XX PN WO9501427-A1.  
 XX PD 12-JAN-1995.  
 XX 28-JUN-1994; 94WO-US007278.  
 XX 02-JUL-1993; 93US-00087163.  
 XX (NEWB-) NEW ENGLAND DEACONESS HOSPITAL.  
 XX PI Liu J, Gurewich V;  
 XX WPI; 1995-060991/08.  
 XX PT Pro-urokinase mutants - have thrombolytic activity but reduced  
 PT fibrinogenolysis activity and non-specific plasminogen activation.  
 XX Claim 15; Fig 1; 46pp; English.

```

XX AAR62991 is the wild type pro-urokinase, from which the new mutants
CC described in AAR62992-R63008 were derived. These mutants retain the
CC thrombolytic activity of the wild type protein, useful for the treatment
CC of thromboembolism, but have a reduced fibrinogenolysis activity and non-
CC specific plasminogen activation. The mutants can therefore be used for
CC the lysis of fibrin clots without inducing systemic bleeding, as can be
CC the case with the wild type protein. (Updated on 25-MAR-2003 to correct
CC PN field.)
XX
SQ Sequence 411 AA;
Query Match 100.0%; Score 793; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 8.7e-54;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCDCLNGGTCVSNKYPSNIHWCNCPKFGQHCCEIDKSKTCYEGNGHFRG 60
DB 1 SNELHQPNSCDCLNGGTCVSNKYPSNIHWCNCPKFGQHCCEIDKSKTCYEGNGHFRG 60
QY 61 KASTDTMGRCPLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVOVGLK 120
DB 61 KASTDTMGRCPLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVOVGLK 120
QY 121 PLVQECMVHDCADGK 135
DB 121 PLVQECMVHDCADGK 135
RESULT 19
AAR63006
ID AAR63006 standard; protein; 411 AA.
XX
AC AAR63006;
XX
DT 25-MAR-2003 (revised)
DT 21-SEP-1995 (first entry)
XX
DE Pro-urokinase mutant Ser175 His197 His300 Ala301 Ala313.
XX
KW Pro-urokinase; thrombolysis; fibrin clot lysis; reduced fibrinogenolysis;
KW non-specific plasminogen activation; systemic bleeding;
KW mutant Ser175 His187 His300 Ala301 Ala313.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Disulfide-bond 11..19
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Domain /note="flexible loop"
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX
PN WO9501427-A1.
XX
PD 12-JAN-1995.
XX
PF 28-JUN-1994; 94WO-US007278.
XX
PR 02-JUL-1993; 93US-00087163.
XX
PA (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.
XX
PI Liu J, Gurewicz V;

```

```

XX WPI; 1995-060991/08.
XX
PT Pro-urokinase mutants - have thrombolytic activity but reduced
PT fibrinogenolysis activity and non-specific plasminogen activation.
XX
PS Claim 16; Fig 1; 46pp; English.
XX
CC AAR62991 is the wild type pro-urokinase, from which the new mutants
CC described in AAR62992-R63008 were derived. These mutants retain the
CC thrombolytic activity of the wild type protein, useful for the treatment
CC of thromboembolism, but have a reduced fibrinogenolysis activity and non-
CC specific plasminogen activation. The mutants can therefore be used for
CC the lysis of fibrin clots without inducing systemic bleeding, as can be
CC the case with the wild type protein. (Updated on 25-MAR-2003 to correct
CC PN field.)
XX
SQ Sequence 411 AA;
Query Match 100.0%; Score 793; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 8.7e-54;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCDCLNGGTCVSNKYPSNIHWCNCPKFGQHCCEIDKSKTCYEGNGHFRG 60
DB 1 SNELHQPNSCDCLNGGTCVSNKYPSNIHWCNCPKFGQHCCEIDKSKTCYEGNGHFRG 60
QY 61 KASTDTMGRCPLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVOVGLK 120
DB 61 KASTDTMGRCPLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVOVGLK 120
QY 121 PLVQECMVHDCADGK 135
DB 121 PLVQECMVHDCADGK 135
RESULT 20
AAR62992
ID AAR62992 standard; protein; 411 AA.
XX
AC AAR62992;
XX
DT 25-MAR-2003 (revised)
DT 21-SEP-1995 (first entry)
XX
DE Pro-urokinase mutant Ala300.
XX
KW Pro-urokinase; thrombolysis; fibrin clot lysis; mutant Ala300;
KW reduced fibrinogenolysis; non-specific plasminogen activation;
KW systemic bleeding.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Disulfide-bond 11..19
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Domain /note="flexible loop"
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX
PN WO9501427-A1.
XX
PD 12-JAN-1995.
XX

```

FT Disulfide-bond 325. .341  
XX Disulfide-bond 352. .380  
XX WO9501427-A1.  
XX  
XX 12-JAN-1995.  
XX  
XX 28-JUN-1994; 94WO-US007278.  
XX  
XX 02-JUL-1993; 93US-00087163.  
XX  
XX (NEWB-) NEW ENGLAND DEACONESS HOSPITAL.  
XX  
XX Liu J, Gurewich V;  
XX WPI; 1995-060991/08.  
XX  
XX Pro-urokinase mutants - have thrombolytic activity but reduced  
XX fibrinogenolysis activity and non-specific plasminogen activation.  
XX  
XX Claim 15; Fig 1; 46pp; English.  
XX  
XX AAR62991 is the wild type pro-urokinase, from which the new mutants  
XX described in AAR62992-R63008 were derived. These mutants retain the  
XX thrombolytic activity of the wild type protein, useful for the treatment  
XX of thromboembolism, but have a reduced fibrinogenolysis activity and non-  
XX specific plasminogen activation. The mutants can therefore be used for  
XX the lysis of fibrin clots without inducing systemic bleeding, as can be  
XX the case with the wild type protein. (Updated on 25-MAR-2003 to correct  
XX PN field.)  
XX  
XX Sequence 411 AA;  
SQ  
Query Match 100.0%; Score 793; DB 2; Length 411;  
Best Local Similarity 100.0%; Pred. No. 8.7e-54; Indels 0; Gaps 0  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0  
QY 1 SNELHQPSPNCCLNGTCTVSNKYFNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60  
DB 1 SNELHQPSPNCCLNGTCTVSNKYFNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60  
QY 61 KASTDTMGPRCLPWNATVLOQTYHAHRSDALQLGLGKHYCRNPNRRRPPWCYVOVGLK 120  
DB 61 KASTDTMGPRCLPWNATVLOQTYHAHRSDALQLGLGKHYCRNPNRRRPPWCYVOVGLK 120  
QY 121 PLVQECMVHDCADGK 135  
DB 121 PLVQECMVHDCADGK 135  
RESULT 22  
AAR63007  
ID AAR63007 standard; protein; 411 AA.  
XX  
XX AAR63007;  
XX  
XX 25-MAR-2003 (revised)  
DT 21-SEP-1995 (first entry)  
XX  
XX Pro-urokinase mutant Ser175 His187 Ala300 Ala301 His313.  
XX  
XX Pro-urokinase; thrombolysis; fibrin clot lysis; reduced fibrinogenolysis;  
XX non-specific plasminogen activation; systemic bleeding;  
XX mutant Ser175 His187 Ala300 Ala301 His313.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH Disulfide-bond 11. .19  
FT Disulfide-bond 13. .31  
FT Disulfide-bond 33. .42  
FT Disulfide-bond 50. .131  
FT Disulfide-bond 102. .126  
FT Disulfide-bond 148. .205  
FT Disulfide-bond 189. .205  
FT Disulfide-bond 197. .268  
FT Disulfide-bond 293. .362  
FT Domain 297. .313  
/note="flexible loop"

PF 28-JUN-1994; 94WO-US007278.  
XX  
XX 02-JUL-1993; 93US-00087163.  
XX  
XX (NEWB-) NEW ENGLAND DEACONESS HOSPITAL.  
XX  
XX Liu J, Gurewich V;  
XX WPI; 1995-060991/08.  
XX  
XX Pro-urokinase mutants - have thrombolytic activity but reduced  
XX fibrinogenolysis activity and non-specific plasminogen activation.  
XX  
XX Claim 5; Fig 1; 46pp; English.  
XX  
XX AAR62991 is the wild type pro-urokinase, from which the new mutants  
XX described in AAR62992-R63008 were derived. These mutants retain the  
XX thrombolytic activity of the wild type protein, useful for the treatment  
XX of thromboembolism, but have a reduced fibrinogenolysis activity and non-  
XX specific plasminogen activation. The mutants can therefore be used for  
XX the lysis of fibrin clots without inducing systemic bleeding, as can be  
XX the case with the wild type protein. (Updated on 25-MAR-2003 to correct  
XX PN field.)  
XX  
XX Sequence 411 AA;  
SQ  
Query Match 100.0%; Score 793; DB 2; Length 411;  
Best Local Similarity 100.0%; Pred. No. 8.7e-54; Indels 0; Gaps 0  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0  
QY 1 SNELHQPSPNCCLNGTCTVSNKYFNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60  
DB 1 SNELHQPSPNCCLNGTCTVSNKYFNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60  
QY 61 KASTDTMGPRCLPWNATVLOQTYHAHRSDALQLGLGKHYCRNPNRRRPPWCYVOVGLK 120  
DB 61 KASTDTMGPRCLPWNATVLOQTYHAHRSDALQLGLGKHYCRNPNRRRPPWCYVOVGLK 120  
QY 121 PLVQECMVHDCADGK 135  
DB 121 PLVQECMVHDCADGK 135  
RESULT 21  
AAR63002  
ID AAR63002 standard; protein; 411 AA.  
XX  
XX AAR63002;  
XX  
XX 25-MAR-2003 (revised)  
DT 21-SEP-1995 (first entry)  
XX  
XX Pro-urokinase mutant Ser175 His187 Gly306.  
XX  
XX Pro-urokinase; thrombolysis; fibrin clot lysis; reduced fibrinogenolysis;  
XX non-specific plasminogen activation; systemic bleeding;  
XX mutant Ser175 His187 Gly306.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH Disulfide-bond 11. .19  
FT Disulfide-bond 13. .31  
FT Disulfide-bond 33. .42  
FT Disulfide-bond 50. .131  
FT Disulfide-bond 71. .113  
FT Disulfide-bond 102. .126  
FT Disulfide-bond 148. .205  
FT Disulfide-bond 189. .205  
FT Disulfide-bond 197. .268  
FT Disulfide-bond 293. .362  
FT Domain 297. .313  
/note="flexible loop"

FT Disulfide-bond 102. .126  
FT Disulfide-bond 148. .279  
FT Disulfide-bond 189. .205  
FT Disulfide-bond 197. .268  
FT Disulfide-bond 293. .362  
FT Domain 297. .313  
FT /note= "flexible loop"  
FT Disulfide-bond 325. .341  
FT Disulfide-bond 352. .380  
XX WO9501427-Al.  
XX  
PD 12-JAN-1995.  
XX  
PF 28-JUN-1994; 94WO-US007278.  
XX  
PR 02-JUL-1993; 93US-00087163.  
XX  
XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.  
XX  
PI Liu J, Gurewich V;  
XX  
DR WPI; 1995-060991/08.  
XX  
XX Pro-urokinase mutants - have thrombolytic activity but reduced  
XX fibrinogenolysis activity and non-specific plasminogen activation.  
XX  
PS Claim 16; Fig 1; 46pp; English.  
XX  
CC AAR62991 is the wild type pro-urokinase, from which the new mutants  
CC described in AAR62992-R63008 were derived. These mutants retain the  
CC thrombolytic activity of the wild type protein, useful for the treatment  
CC of thromboembolism, but have a reduced fibrinogenolysis activity and non-  
CC specific plasminogen activation. The mutants can therefore be used for  
CC the lysis of fibrin clots without inducing systemic bleeding, as can be  
CC the case with the wild type protein. (Updated on 25-MAR-2003 to correct  
CC PN field.)  
XX  
SQ Sequence 411 AA;  
  
Query Match 100.0%; Score 793; DB 2; Length 411;  
Best Local Similarity 100.0%; Pred. No. 8.7e-54;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKFGGQHCIDKSKTCYEGNGHFYRG 60  
DB 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKFGGQHCIDKSKTCYEGNGHFYRG 60  
  
QY 61 KASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRRRRCWYQVGLK 120  
DB 61 KASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRRRRCWYQVGLK 120  
  
QY 121 PLVQECMVHDCADGK 135  
DB 121 PLVQECMVHDCADGK 135  
  
RESULT 23  
AAR62995  
ID AAR62995 standard; protein; 411 AA.  
XX  
AC AAR62995;  
XX  
DT 25-MAR-2003 (revised)  
DT 21-SEP-1995 (first entry)  
XX  
DE Pro-urokinase mutant Ala300 Ala301.  
XX  
KW Pro-urokinase; thrombolysis; fibrin clot lysis; reduced fibrinogenolysis;  
KW non-specific plasminogen activation; mutant Ala300 Ala301;  
KW systemic bleeding.  
XX  
OS Homo sapiens.

XX Key Location/Qualifiers  
FH Disulfide-bond 11. .19  
FT Disulfide-bond 13. .31  
FT Disulfide-bond 33. .42  
FT Disulfide-bond 50. .131  
FT Disulfide-bond 71. .113  
FT Disulfide-bond 102. .126  
FT Disulfide-bond 148. .279  
FT Disulfide-bond 189. .205  
FT Disulfide-bond 197. .268  
FT Disulfide-bond 293. .362  
FT Domain 297. .313  
FT /note= "flexible loop"  
FT Disulfide-bond 325. .341  
FT Disulfide-bond 352. .380  
XX WO9501427-Al.  
XX  
PD 12-JAN-1995.  
XX  
PF 28-JUN-1994; 94WO-US007278.  
XX  
PR 02-JUL-1993; 93US-00087163.  
XX  
XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.  
XX  
PI Liu J, Gurewich V;  
XX  
DR WPI; 1995-060991/08.  
XX  
XX Pro-urokinase mutants - have thrombolytic activity but reduced  
XX fibrinogenolysis activity and non-specific plasminogen activation.  
XX  
PS Claim 9; Fig 1; 46pp; English.  
XX  
CC AAR62991 is the wild type pro-urokinase, from which the new mutants  
CC described in AAR62992-R63008 were derived. These mutants retain the  
CC thrombolytic activity of the wild type protein, useful for the treatment  
CC of thromboembolism, but have a reduced fibrinogenolysis activity and non-  
CC specific plasminogen activation. The mutants can therefore be used for  
CC the lysis of fibrin clots without inducing systemic bleeding, as can be  
CC the case with the wild type protein. (Updated on 25-MAR-2003 to correct  
CC PN field.)  
XX  
SQ Sequence 411 AA;  
  
Query Match 100.0%; Score 793; DB 2; Length 411;  
Best Local Similarity 100.0%; Pred. No. 8.7e-54;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKFGGQHCIDKSKTCYEGNGHFYRG 60  
DB 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKFGGQHCIDKSKTCYEGNGHFYRG 60  
  
QY 61 KASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRRRRCWYQVGLK 120  
DB 61 KASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRRRRCWYQVGLK 120  
  
QY 121 PLVQECMVHDCADGK 135  
DB 121 PLVQECMVHDCADGK 135  
  
RESULT 24  
AAR63000  
ID AAR63000 standard; protein; 411 AA.  
XX  
AC AAR63000;  
XX  
DT 25-MAR-2003 (revised)  
DT 21-SEP-1995 (first entry)  
XX



QY	121	PLVQECMVHDCADGK 135	
Db	121	PLVQECMVHDCADGK 135	
RESULT 26			
AAR62996			
ID	AAR62996	standard; protein; 411 AA.	
XX			
AC	AAR62996;		
XX			
DT	25-MAR-2003	(revised)	
DT	21-SEP-1995	(first entry)	
XX			
DE	Pro-urokinase mutant His300 Ala301.		
XX			
KW	Pro-urokinase; thrombolysis; fibrin clot lysis; reduced fibrinogenolysis;		
KW	non-specific plasminogen activation; mutant His300 Ala301;		
KW	systemic bleeding.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	Disulfide-bond	11. .19	
FT	Disulfide-bond	13. .31	
FT	Disulfide-bond	33. .42	
FT	Disulfide-bond	50. .131	
FT	Disulfide-bond	71. .113	
FT	Disulfide-bond	102. .126	
FT	Disulfide-bond	148. .279	
FT	Disulfide-bond	189. .205	
FT	Disulfide-bond	197. .268	
FT	Disulfide-bond	293. .362	
FT	Domain	297. .313	
FT		/note= "flexible loop"	
FT	Disulfide-bond	325. .341	
FT	Disulfide-bond	352. .380	
XX	WO9501427-A1.		
XX			
PD	12-JAN-1995.		
XX			
DF	28-JUN-1994;	94WO-US007278.	
XX			
PR	02-JUL-1993;	93US-00087163.	
XX			
PA	(NEW-) NEW ENGLAND DEACONESS HOSPITAL.		
XX			
PI	Liu J, Gurewich V;		
XX			
DR	WPI; 1995-060991/08.		
XX			
FT	Pro-urokinase mutants - have thrombolytic activity but reduced		
FT	fibrinogenolysis activity and non-specific plasminogen activation.		
XX			
PS	Claim 9; Fig 1; 46pp; English.		
XX			
CC	AAR62991 is the wild type pro-urokinase, from which the new mutants		
CC	described in AAR62992-R63008 were derived. These mutants retain the		
CC	thrombolytic activity of the wild type protein, useful for the treatment		
CC	of thromboembolism, but have a reduced fibrinogenolysis activity and non-		
CC	specific plasminogen activation. The mutants can therefore be used for		
CC	the lysis of fibrin clots without inducing systemic bleeding, as can be		
CC	the case with the wild type protein. (Updated on 25-MAR-2003 to correct		
CC	PN field.)		
XX			
SQ	Sequence 411 AA;		
Query Match			
	Best Local Similarity	100.0%;	Score 793; DB 2; Length 411;
	Matches 135; Conservative	0; Mismatches	0; Indels 0; Gaps 0;
QY	1	SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGQCHCEIDKSKTCYEGNGHFYRG 60	



SQ Sequence 411 AA;

Query Match 100.0%; Score 793; DB 2; Length 411;  
 Best Local Similarity 100.0%; Pred. No. 8.7e-54;  
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYSNIHWNCNPKKFGGQHCETDKSKTCYEGNGHFFYRG 60  
 DB 1 SNELHQVPSNCDCLNGGTCVSNKYSNIHWNCNPKKFGGQHCETDKSKTCYEGNGHFFYRG 60  
 QY 61 KASTDTMGPRCLPWSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRPRPCVYQVGLK 120  
 DB 61 KASTDTMGPRCLPWSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRPRPCVYQVGLK 120  
 QY 121 PLVQECMVHDCADGK 135  
 DB 121 PLVQECMVHDCADGK 135

RESULT 28  
 AAR62991  
 ID AAR62991 standard; protein; 411 AA.  
 XX AC AAR62991;  
 XX DT 25-MAR-2003 (revised)  
 XX DT 21-SEP-1995 (first entry)  
 XX DE Pro-urokinase.  
 XX KW Pro-urokinase; thrombolysis; fibrin clot lysis.  
 XX OS Homo sapiens.  
 XX PH Key Location/Qualifiers  
 FT Disulfide-bond 11..19  
 FT Disulfide-bond 13..31  
 FT Disulfide-bond 33..42  
 FT Disulfide-bond 50..131  
 FT Disulfide-bond 71..113  
 FT Disulfide-bond 102..126  
 FT Disulfide-bond 148..279  
 FT Disulfide-bond 189..205  
 FT Disulfide-bond 197..268  
 FT Disulfide-bond 293..362  
 FT Domain 297..313  
 FT /note= "flexible loop"  
 FT Disulfide-bond 325..341  
 FT Disulfide-bond 352..380  
 XX PN WO9501427-A1.  
 XX PD 12-JAN-1995.  
 XX PF 28-JUN-1994; 94WO-US007278.  
 XX PR 02-JUL-1993; 93US-00087163.  
 XX PA (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.  
 XX PI Liu J, Gurewich V;  
 XX WPI; 1995-060991/08.  
 XX DR Pro-urokinase mutants - have thrombolytic activity but reduced  
 XX FT fibrinogenolysis activity and non-specific plasminogen activation.  
 XX PS Disclosure; Fig 1; 46pp; English.  
 XX CC AAR62991 is the wild type pro-urokinase, from which the new mutants  
 CC described in AAR62992-R3008 were derived. These mutants retain the  
 CC thrombolytic activity of the wild type protein, useful for the treatment  
 CC of thromboembolism, but have a reduced fibrinogenolysis activity and non-

CC specific plasminogen activation. The mutants can therefore be used for  
 CC the lysis of fibrin clots without inducing systemic bleeding, as can be  
 CC the case with the wild type protein. (Updated on 25-MAR-2003 to correct  
 CC PN field.)  
 XX SQ Sequence 411 AA;

Query Match 100.0%; Score 793; DB 2; Length 411;  
 Best Local Similarity 100.0%; Pred. No. 8.7e-54;  
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYSNIHWNCNPKKFGGQHCETDKSKTCYEGNGHFFYRG 60  
 DB 1 SNELHQVPSNCDCLNGGTCVSNKYSNIHWNCNPKKFGGQHCETDKSKTCYEGNGHFFYRG 60  
 QY 61 KASTDTMGPRCLPWSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRPRPCVYQVGLK 120  
 DB 61 KASTDTMGPRCLPWSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRPRPCVYQVGLK 120  
 QY 121 PLVQECMVHDCADGK 135  
 DB 121 PLVQECMVHDCADGK 135

RESULT 29  
 AAR62998  
 ID AAR62998 standard; protein; 411 AA.  
 XX AC AAR62998;  
 XX DT 25-MAR-2003 (revised)  
 XX DT 21-SEP-1995 (first entry)  
 XX DE Pro-urokinase mutant Ala313.  
 XX KW Pro-urokinase; thrombolysis; fibrin clot lysis; mutant Ala313;  
 KW reduced fibrinogenolysis; non-specific plasminogen activation;  
 XX KW systemic bleeding.  
 XX OS Homo sapiens.  
 XX PH Key Location/Qualifiers  
 FT Disulfide-bond 11..19  
 FT Disulfide-bond 13..31  
 FT Disulfide-bond 33..42  
 FT Disulfide-bond 50..131  
 FT Disulfide-bond 71..113  
 FT Disulfide-bond 102..126  
 FT Disulfide-bond 148..279  
 FT Disulfide-bond 189..205  
 FT Disulfide-bond 197..268  
 FT Disulfide-bond 293..362  
 FT Domain 297..313  
 FT /note= "flexible loop"  
 FT Disulfide-bond 325..341  
 FT Disulfide-bond 352..380  
 XX PN WO9501427-A1.  
 XX PD 12-JAN-1995.  
 XX PF 28-JUN-1994; 94WO-US007278.  
 XX PR 02-JUL-1993; 93US-00087163.  
 XX PA (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.  
 XX PI Liu J, Gurewich V;  
 XX WPI; 1995-060991/08.  
 XX DR Pro-urokinase mutants - have thrombolytic activity but reduced  
 XX FT fibrinogenolysis activity and non-specific plasminogen activation.  
 XX PS Disclosure; Fig 1; 46pp; English.  
 XX CC AAR62991 is the wild type pro-urokinase, from which the new mutants  
 CC described in AAR62992-R3008 were derived. These mutants retain the  
 CC thrombolytic activity of the wild type protein, useful for the treatment  
 CC of thromboembolism, but have a reduced fibrinogenolysis activity and non-

```
XX PS Claim 11; Fig 1; 46pp; English.
XX CC
XX CC AAR62991 is the wild type pro-urokinase, from which the new mutants
XX CC described in AAR62992-R63008 were derived. These mutants retain the
XX CC thrombolytic activity of the wild type protein, useful for the treatment
XX CC of thromboembolism, but have a reduced fibrinogenolysis activity and non-
XX CC specific plasminogen activation. The mutants can therefore be used for
XX CC the lysis of fibrin clots without inducing systemic bleeding, as can be
XX CC the case with the wild type protein. (Updated on 25-MAR-2003 to correct
XX CC PN field.)
XX SQ Sequence 411 AA;
XX Query Match 100.0%; Score 793; DB 2; Length 411;
XX Best Local Similarity 100.0%; Pred. No. 8.7e-54;
XX Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 SNELHQPVSNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
XX DB 1 SNELHQPVSNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
XX QY 61 KASDTMTGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRPNRRPWCYVQVGLK 120
XX DB 61 KASDTMTGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRPNRRPWCYVQVGLK 120
XX QY 121 PLVQECMVHDCADGK 135
XX DB 121 PLVQECMVHDCADGK 135
XX RESULT 30
XX AAR62993
XX ID AAR62993 standard; protein; 411 AA.
XX AC AAR62993;
XX DT 25-MAR-2003 (revised)
XX DT 21-SEP-1995 (first entry)
XX DE Pro-urokinase mutant His300.
XX KW Pro-urokinase; thrombolysis; fibrin clot lysis; mutant His300;
XX KW reduced fibrinogenolysis; non-specific plasminogen activation;
XX KW systemic bleeding.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Disulfide-bond 11..19
XX FT Disulfide-bond 13..31
XX FT Disulfide-bond 33..42
XX FT Disulfide-bond 50..131
XX FT Disulfide-bond 71..113
XX FT Disulfide-bond 102..126
XX FT Disulfide-bond 148..279
XX FT Disulfide-bond 189..205
XX FT Disulfide-bond 197..268
XX FT Disulfide-bond 293..362
XX FT Domain 297..313
XX FT /note= "flexible loop"
XX FT Disulfide-bond 325..341
XX FT Disulfide-bond 352..380
XX PN WO9501427-A1.
XX PD 12-JAN-1995.
XX PF 28-JUN-1994; 94WO-US007278.
XX PR 02-JUL-1993; 93US-00087163.
XX PA (NEW-) NEW ENGLAND DEACONESS HOSPITAL.

XX PS Liu J, Gurewich V;
XX CC
XX CC WPI; 1995-060991/08.
XX CC
XX CC Pro-urokinase mutants - have thrombolytic activity but reduced
XX CC fibrinogenolysis activity and non-specific plasminogen activation.
XX CC
XX CC Claim 5; Fig 1; 46pp; English.
XX CC
XX CC AAR62991 is the wild type pro-urokinase, from which the new mutants
XX CC described in AAR62992-R63008 were derived. These mutants retain the
XX CC thrombolytic activity of the wild type protein, useful for the treatment
XX CC of thromboembolism, but have a reduced fibrinogenolysis activity and non-
XX CC specific plasminogen activation. The mutants can therefore be used for
XX CC the lysis of fibrin clots without inducing systemic bleeding, as can be
XX CC the case with the wild type protein. (Updated on 25-MAR-2003 to correct
XX CC PN field.)
XX SQ Sequence 411 AA;
XX Query Match 100.0%; Score 793; DB 2; Length 411;
XX Best Local Similarity 100.0%; Pred. No. 8.7e-54;
XX Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 SNELHQPVSNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
XX DB 1 SNELHQPVSNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
XX QY 61 KASDTMTGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRPNRRPWCYVQVGLK 120
XX DB 61 KASDTMTGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRPNRRPWCYVQVGLK 120
XX QY 121 PLVQECMVHDCADGK 135
XX DB 121 PLVQECMVHDCADGK 135
XX RESULT 31
XX AAR92926
XX ID AAR92926 standard; protein; 411 AA.
XX AC AAR92926;
XX DT 03-AUG-1996 (first entry)
XX DE Pro-urokinase.
XX KW Pro-urokinase; plasminogen activator; fusion drug; drug delivery;
XX KW platelet; cardiovascular disease; thrombolytic.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Region 1..132
XX FT /label= A-chain
XX FT Domain 1..45
XX FT /label= Growth_factor_domain
XX FT Disulfide-bond 11
XX FT Disulfide-bond 13
XX FT /note= "disulfide between Cys11 and Cys19"
XX FT Disulfide-bond 33
XX FT /note= "disulfide bond between Cys13 and Cys31"
XX FT Disulfide-bond 33
XX FT /note= "disulfide bond between Cys33 and Cys42"
XX FT Domain 46..132
XX FT /label= Kringle_domain
XX FT Disulfide-bond 50
XX FT /note= "disulfide bond between Cys50 and Cys131"
XX FT Disulfide-bond 71
XX FT /note= "disulfide bond between Cys71 and Cys113"
XX FT Disulfide-bond 102
XX FT /note= "disulfide bond between Cys102 and Cys126"
XX FT Region 133..158
```

```

FT FT Disulfide-bond 148 /label= Linker_region
FT FT 156..157 /note= "disulfide bond between Cys148 and Cys279"
FT FT Cleavage-site 158..159 /note= "thrombin cleavage site"
FT FT Cleavage-site 159..159 /note= "plasmin cleavage site"
FT FT Region 159..411 /label= B-chain
FT FT Disulfide-bond 189 /note= "disulfide bond between Cys189 and Cys205"
FT FT Disulfide-bond 197 /note= "disulfide bond between Cys197 and Cys268"
FT FT Disulfide-bond 293 /note= "disulfide bond between Cys293 and Cys362"
FT FT Disulfide-bond 325 /note= "disulfide bond between Cys325 and Cys341"
FT FT Disulfide-bond 352 /note= "disulfide bond between Cys352 and Cys380"
XX XX WO9604004-A1.
XX XX 15-FEB-1996.
XX XX 03-AUG-1995; 95WO-US009848.
XX XX 05-AUG-1994; 94US-00286748.
XX XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.
XX XX Gurewich V;
XX XX WPI; 1996-129123/13.
XX XX N-PSDB; AAT18237.
XX XX Fusion product of plasminogen activator A chain and drug - targeted to
XX XX platelets, useful for treatment of cardiovascular disease.
XX XX Claim 3; Page 39-40; 61pp; English.
XX XX A portion of the A-chain (pref. amino acids 1-132) of pro-urokinase
XX XX (AAR92926) can be linked to a drug, e.g. hirudin, somatostatin or their
XX XX analogues (see also AAR92925 and AAR92927-33) and used in the prodn. of
XX XX new fusion drugs. The constructs can be obt'd. by expression of the
XX XX appropriate nucleotide sequences in transformed host cells. When
XX XX administered to a patient, the A-chain binds the fusion drug to the
XX XX platelet outer membrane, i.e. to the site of thrombosis or vascular
XX XX injury. Cleavage sites for thrombin and/or plasmin with the fusion drug
XX XX allow the release of the drug at the target site
XX XX Sequence 411 AA;
XX XX
XX XX Query Match 100.0%; Score 793; DB 2; Length 411;
XX XX Best Local Similarity 100.0%; Pred. No. 8.7e-54;
XX XX Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPVSNDCLNGGTCVSNKVFNSIHWNCNPKFGQGCETDKSKTCYEGNGHPIRG 60
DB 1 SNELHQPVSNDCLNGGTCVSNKVFNSIHWNCNPKFGQGCETDKSKTCYEGNGHPIRG 60
QY 61 KASDTDMGRPCLPWNATVLOQTYHAHRSALQGLGKHNYCRPNDRRRPWCYVOVGLK 120
DB 61 KASDTDMGRPCLPWNATVLOQTYHAHRSALQGLGKHNYCRPNDRRRPWCYVOVGLK 120
QY 121 FLVQECMVHDCADGK 135
DB 121 FLVQECMVHDCADGK 135
XX XX
XX XX RESULT 32
XX XX AAY92836
XX XX ID AAY92836 standard; protein; 411 AA.

```

```

AC AAY92836;
XX XX
XX XX 29-AUG-2000 (first entry)
XX XX
XX XX Urokinase plasminogen activator (uPA).
XX XX
XX XX N-terminal; pro-uPA; urokinase plasminogen activator; receptor; uPAR;
XX XX anti-cancer; anti-metastatic; anti-proliferative; anti-atherosclerotic;
XX XX anti-thrombotic; anti-angiogenic; anti-inflammatory; anti-arthritis;
XX XX anti-fibrotic; apoptotic; vasotropic; anti-diabetic; ophthalmological;
XX XX thrombolytic.
XX XX
XX XX Homo sapiens.
XX XX
XX XX Key Location/Qualifiers
XX XX Disulfide-bond 11..19
XX XX Disulfide-bond 13..31
XX XX Disulfide-bond 33..42
XX XX Disulfide-bond 50..131
XX XX Disulfide-bond 71..113
XX XX Disulfide-bond 102..126
XX XX Disulfide-bond 148..279
XX XX Disulfide-bond 189..205
XX XX Disulfide-bond 197..268
XX XX Disulfide-bond 293..362
XX XX Disulfide-bond 325..341
XX XX Disulfide-bond 352..380
XX XX
XX XX WO200026353-A1.
XX XX
XX XX 11-MAY-2000.
XX XX
XX XX 28-OCT-1999; 99WO-US025210.
XX XX
XX XX 29-OCT-1998; 98US-00181816.
XX XX
XX XX (ANGS-) ANGSTROM PHARM INC.
XX XX
XX XX Mazar AP, Jones TR;
XX XX
XX XX WPI; 2000-365605/31.
XX XX
XX XX New cyclic peptide, useful for treatment or diagnosis of e.g. tumors and
XX XX other diseases involving cell proliferation or migration, targets the
XX XX urokinase plasminogen activator receptor.
XX XX
XX XX Disclosure; Fig 1; 93pp; English.
XX XX
XX XX The present sequence shows the wild-type urokinase plasminogen activator
XX XX (uPA). Cyclic peptides based on the amino acids residues 20-30 (the
XX XX receptor-binding region) of uPA are claimed. These cyclic peptides target
XX XX the uPA receptor (uPAR), allowing therapeutic or diagnostic agents to be
XX XX delivered to uPAR-expressing cells. The cyclic peptides are used,
XX XX optionally when linked to a therapeutic agent, to inhibit migration,
XX XX invasion and proliferation of cells, or angiogenesis, or to induce
XX XX apoptosis. Particularly they are used, in human or veterinary medicine,
XX XX to treat diseases characterized by these processes, e.g. solid tumors,
XX XX leukemia or lymphoma (or their metastases); benign hyperplasia;
XX XX atherosclerosis; restenosis; ischaemia; deep vein thrombosis; neovascular
XX XX glaucoma; diabetic retinopathy; arthritis; fibrosis; bone fracture etc.,
XX XX most particularly growth, invasion and metastasis of tumors. When
XX XX labeled, the cyclic peptides can be used for diagnostic detection of uPAR
XX XX (a marker of metastasis) on cells, tissues etc., in vivo or in vitro, and
XX XX when immobilized they are used to isolate uPAR or cells that express
XX XX them. The cyclic peptides are stable, soluble in water, bind strongly to
XX XX uPAR, are relatively inexpensive to produce and may be derivatized by
XX XX attachment of therapeutic or diagnostic agents without significantly
XX XX affecting their binding. Since they target uPAR, they should have
XX XX relatively low systemic toxicity and only low doses are required
XX XX
XX XX Sequence 411 AA;
XX XX
XX XX Query Match 100.0%; Score 793; DB 3; Length 411;

```

Best Local Similarity 100.0%; Pred. No. 8.7e-54; Mismatches 0; Indels 0; Gaps 0;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60  
DB 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60  
QY 61 KASDTDTMGRPCLPWSNATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 120  
DB 61 KASDTDTMGRPCLPWSNATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 120  
QY 121 PLVQECMVHDCADGK 135  
DB 121 PLVQECMVHDCADGK 135

RESULT 33  
AAE16544  
ID AAE16544 standard; protein; 411 AA.  
XX AC AAE16544;  
XX DT 09-APR-2002 (first entry)  
XX DE Human urokinase-type plasminogen activator tcupA and scuPA protein.  
XX KW Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;  
XX KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;  
XX KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;  
XX KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;  
XX KW clotting disorder; uterine contraction disorder; respiratory disease;  
XX KW male impotence; adult respiratory distress syndrome; tcupA; scuPA;  
XX KW two chain urokinase; single chain urokinase.  
XX OS Homo sapiens.  
XX PN WO200197752-A2.  
XX PD 27-DEC-2001.  
XX PF 13-JUN-2001; 2001WO-US018976.  
XX PR 20-JUN-2000; 2000US-0212874P.  
XX PA (UYPE-) UNIV PENNSYLVANIA.  
XX PI Cines DB, Higazi AA;  
XX DR WPI; 2002-122240/16.  
XX DR N-PSDB; AAD27077.  
XX PT Composition for modulating muscle cell and tissue contractility for  
XX PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,  
XX PT comprising domains from urokinase-type plasminogen activator.  
XX PS Claim 9; Fig 1C; 117pp; English.  
XX CC The invention relates to a composition comprising one or more domains of  
XX CC urokinase-type plasminogen activator (uPA). The composition is used to  
XX CC modulate the contractility and angiogenic activity of a mammalian muscle,  
XX CC endothelial cell or tissue. The composition is used for treating stroke,  
XX CC hypotension, hypertension, atherosclerosis, heart attack, microvascular  
XX CC occlusions, thrombotic microangiopathies, surgically induced thrombotic  
XX CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell  
XX CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,  
XX CC diabetic retinopathy, wound healing, clotting disorder, uterine  
XX CC contraction disorder, male impotence, respiratory disease or condition  
XX CC such as asthma, adult respiratory distress syndrome, primary pulmonary  
XX CC hypertension, microvascular thrombotic occlusion, and a disorder  
XX CC associated with chronic intrapulmonary fibrin formation. The present  
XX CC sequence is human urokinase-type plasminogen activator (uPA) two chain  
XX CC urokinase (tcuPA) and single chain urokinase (scuPA) protein

SQ Sequence 411 AA;  
Query Match 100.0%; Score 793; DB 5; Length 411;  
Best Local Similarity 100.0%; Pred. No. 8.7e-54;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60  
DB 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60  
QY 61 KASDTDTMGRPCLPWSNATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 120  
DB 61 KASDTDTMGRPCLPWSNATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 120  
QY 121 PLVQECMVHDCADGK 135  
DB 121 PLVQECMVHDCADGK 135

RESULT 34  
AAR10334  
ID AAR10334 standard; protein; 412 AA.  
XX AC AAR10334;  
XX DT 25-MAR-2003 (revised)  
XX DT 05-APR-1991 (first entry)  
XX DE Recombinant single-chain urine plasminogen activator.  
XX KW Single chain urine plasminogen activator; enterobacteriaceae; scu-PA;  
XX KW cardiac infarction; pulmonary embolism; arterial occlusive disease.  
XX OS Synthetic.  
XX PN BP408945-A.  
XX PD 23-JAN-1991.  
XX PF 29-JUN-1990; 90EP-00112400.  
XX PR 19-JUL-1989; 89DE-03923866.  
XX PA (CHEF) GRUENENTHAL GMBH.  
XX PI Brigelius R, Flohe L, Hillen W, Steffens GJ, Strassburg W;  
XX PI Wilhelm M;  
XX DR WPI; 1991-023590/04.  
XX DR N-PSDB; AAQ10282.  
XX PT New plasmids for expressing urine plasminogen activator - in  
XX PT enterobacteriaceae have operon contg. promoter, ribosome binding site,  
XX PT synthetic gene and terminator, etc.  
XX PS Disclosure; Fig 15; 44pp; German.  
XX CC This recombinant single-chain urine-plasminogen activator (scu-PA), is  
XX CC encoded by a structural gene inserted into plasmid pBR322 (from which the  
XX CC ntc/boam and/or tet resistance genes have been removed), following and a  
XX CC insertion of a multiple cloning site (see AAQ10281), a terminator and a  
XX CC synthetic trp promoter. The resultant expression construct can be used to  
XX CC produce high yields of this scu-PA protein which can be refolded to  
XX CC therapeutically active recombinant scu-PA. This is used in the treatment  
XX CC of e.g. cardiac infarction, pulmonary embolism or arterial occlusive  
XX CC disease. (Updated on 25-MAR-2003 to correct PA field.)  
XX SQ Sequence 412 AA;  
Query Match 100.0%; Score 793; DB 2; Length 412;  
Best Local Similarity 100.0%; Pred. No. 8.7e-54;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWNCNPKKFGQHQCEIDKSKTCYEGNGHFYRG 60  
 Db 2 SNELHQPNSCCLNGGTCVSNKYFSNIHWNCNPKKFGQHQCEIDKSKTCYEGNGHFYRG 61  
 QY 61 KASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120  
 Db 62 KASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 121  
 QY 121 PLVQECMVHDCADGK 135  
 Db 122 PLVQECMVHDCADGK 136

## RESULT 35

AAW24579  
 ID AAW24579 standard; protein; 424 AA.

XX AC AAW24579;  
 XX AC

DT 25-MAR-2003 (revised)  
 DT 11-NOV-1997 (first entry)

XX Inhibitor resistant modified urokinase.

XX Urokinase; inhibitor resistant; plasminogen; human; whey acid protein;  
 KW plasminogen activator inhibitor-1; plasminogen activator; protease; WAP;  
 KW plasmin; proteolytic enzyme; clot lysis; heart attack; heart muscle.

XX Homo sapiens.

OS Synthetic.

XX Key  
 FH Peptide  
 FT Location/Qualifiers  
 FT 1..19  
 FT /note= "whey acid protein signal peptide"  
 FT Protein  
 FT 20..430  
 FT /note= "urokinase"  
 FT Domain  
 FT 20..64  
 FT /note= "E-domain"  
 FT Domain  
 FT 61..150  
 FT /note= "Kringle-1 domain"  
 FT Domain  
 FT 179..424  
 FT /note= "P-domain"  
 FT Misc-difference 197..198  
 FT /note= "Site of 6 residue (Arg-His-Arg-Gly-Ser)  
 FT deletion"

XX US5648253-A.

XX 15-JUL-1997.

XX 08-SEP-1992; 92US-00942157.

XX 20-DEC-1990; 90US-00631673.

XX (TSIT-) TSI CORP.

XX Wei C;

XX WPI; 1997-372062/34.

XX N-PSDB; AAT80076.

XX Deletion-modified urokinase protein - with increased resistance to  
 PT inhibition by plasminogen activator inhibitor-1.

XX Claim 2; Col; 16pp; English.

XX This sequence represents the modified full length urokinase of the  
 CC invention, including the whey acid protein (WAP) signal peptide. This  
 CC sequence had residues 179-184 of the wild type urokinase sequence (see  
 CC AAW24578) deleted. This modified urokinase cleaves plasminogen, and has a  
 CC lower binding affinity for plasminogen activator inhibitor-1 than the  
 CC corresponding unmodified urokinase. Urokinase is one of two types of  
 CC mammalian plasminogen activators (PA), the other being tissue type PA.

CC PAs catalyze the conversion of the circulating zymogen plasminogen to the  
 CC broad spectrum protease plasmin by limited proteolysis. The modified  
 CC urokinase can be used for clot lysis, specifically to dissolve heart  
 CC attack-causing clots before they cause permanent damage to heart muscle.  
 CC The urokinase mutant is more resistant to inhibition by plasminogen  
 CC activator inhibitors than the unmodified urokinase. It can be selectively  
 CC expressed and secreted from the mammary glands of transgenic animals.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX

SQ Sequence 424 AA;

Query Match 100.0%; Score 793; DB 2; Length 424;

Best Local Similarity 100.0%; Pred. No. 9e-54;

Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWNCNPKKFGQHQCEIDKSKTCYEGNGHFYRG 60

Db 20 SNELHQPNSCCLNGGTCVSNKYFSNIHWNCNPKKFGQHQCEIDKSKTCYEGNGHFYRG 79

QY 61 KASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120

Db 80 KASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 135

QY 121 PLVQECMVHDCADGK 135

Db 140 PLVQECMVHDCADGK 154

## RESULT 36

AAW24578

ID AAW24578 standard; protein; 430 AA.

XX AC AAW24578;

XX 25-MAR-2003 (revised)

DT 11-NOV-1997 (first entry)

XX Inhibitor resistant urokinase.

XX Urokinase; inhibitor resistant; plasminogen; human; whey acid protein;  
 KW plasminogen activator inhibitor-1; plasminogen activator; protease; WAP;  
 KW plasmin; proteolytic enzyme; clot lysis; heart attack; heart muscle.

XX Homo sapiens.

OS Synthetic.

XX Key  
 FH Peptide  
 FT Location/Qualifiers  
 FT 1..19  
 FT /note= "whey acid protein signal peptide"  
 FT Protein  
 FT 20..430  
 FT /note= "urokinase"  
 FT Domain  
 FT 20..64  
 FT /note= "E-domain"  
 FT Domain  
 FT 61..150  
 FT /note= "Kringle-1 domain"  
 FT Domain  
 FT 179..430  
 FT /note= "P-domain"  
 FT Misc-difference 198..203  
 FT /note= "deleted in modified urokinase of the invention"

XX US5648253-A.

XX 15-JUL-1997.

XX 08-SEP-1992; 92US-00942157.

XX 20-DEC-1990; 90US-00631673.

XX (TSIT-) TSI CORP.

XX Wei C;

XX WPI; 1997-372062/34.

DR N-PSDB; AAT80075.  
 XX Deletion-modified urokinase protein - with increased resistance to  
 PT inhibition by plasminogen activator inhibitor-1.  
 XX  
 XX PS Disclosure; Col 15-18; 16pp; English.  
 XX  
 XX This sequence represents the full length urokinase, including the whey  
 CC acid protein (WAP) signal peptide. This sequence has residues 179-184 of  
 CC the urokinase sequence deleted to create the modified urokinase of the  
 CC invention. The modified urokinase (see AAW24579) cleaves plasminogen, and  
 CC has a lower binding affinity for plasminogen activator inhibitor-1 than  
 CC the corresponding unmodified urokinase. Urokinase is one of two types of  
 CC mammalian plasminogen activators (PA), the other being tissue type PA.  
 CC PAs catalyze the conversion of the circulating zymogen plasminogen to the  
 CC broad spectrum protease plasmin by limited proteolysis. The modified  
 CC urokinase can be used for clot lysis, specifically to dissolve heart  
 CC attack-causing clots before they cause permanent damage to heart muscle.  
 CC The urokinase mutant is more resistant to inhibition by plasminogen  
 CC activator inhibitors than the unmodified urokinase. It can be selectively  
 CC expressed and secreted from the mammary glands of transgenic animals.  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 XX  
 XX SQ Sequence 430 AA;  
 Query Match 100.0%; Score 793; DB 2; Length 430;  
 Best Local Similarity 100.0%; Pred. No. 9.1e-54;  
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
 Db 20 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCIDKSKTCYEGNGHFYRG 79  
 QY 61 KASDTMTGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120  
 Db 80 KASDTMTGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRPNDRRRPWCYVQVGLK 139  
 QY 121 PLVQECMVHDCADGK 135  
 Db 140 PLVQECMVHDCADGK 154  
 RESULT 37  
 AAP50114  
 ID AAP50114 standard; protein; 431 AA.  
 XX  
 XX AC AAP50114;  
 XX  
 XX DT 27-SEP-1991 (first entry)  
 XX  
 XX DE Sequence encoded by the signal sequence and noncoding region of the pro-  
 XX UK structural gene (sequence II).  
 XX  
 XX KW Enzyme; thrombosis therapy; embolic disease; single-chain pro-urokinase.  
 XX  
 XX OS Homo sapiens.  
 XX  
 XX FH Key Location/Qualifiers  
 XX FT Peptide 1..20  
 XX FT /label= signal peptide  
 XX FT Domain 21..177  
 XX FT /label= A chain  
 XX FT Domain 179..431  
 XX FT /label= B chain  
 XX  
 XX EP154272-A.  
 XX  
 XX PD 11-SEP-1985.  
 XX  
 XX PF 23-FEB-1985; 85EP-00102031.  
 XX  
 XX PR 27-FEB-1984; 84JP-00037119.  
 XX 31-JAN-1985; 85JP-00017969.

XX (GREC ) GREEN CROSS CORP.  
 XX Hiramatsu R, Kaneda T, Nagai M, Rimura H, Nishida M, Suyama T;  
 XX  
 XX DR WPI; 1985-224693/37.  
 XX N-PSDB; AAN50138.  
 XX  
 XX PT Glycosylated single-chain pro-urokinase - prepd. by cultivating animal  
 XX cells transformed by DNA prepd. from m RNA.  
 XX  
 XX PS Disclosure; Page 8-10; 64pp; English.  
 XX  
 XX CC The inventors claim a method of producing single-chain pro-urokinase by  
 CC using as template, mRNA obtd. from cells of an established human kidney-  
 CC derived cell line. The urokinase is used to treat thrombosis and embolic  
 CC diseases as well as in the treatment of diseases in combination with  
 CC anticancer agents  
 XX  
 XX SQ Sequence 431 AA;  
 Query Match 100.0%; Score 793; DB 1; Length 431;  
 Best Local Similarity 100.0%; Pred. No. 9.1e-54;  
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0  
 QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
 Db 21 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80  
 QY 61 KASDTMTGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120  
 Db 81 KASDTMTGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRPNDRRRPWCYVQVGLK 140  
 QY 121 PLVQECMVHDCADGK 135  
 Db 141 PLVQECMVHDCADGK 155  
 RESULT 38  
 AAP60783  
 ID AAP60783 standard; protein; 431 AA.  
 XX  
 XX AC AAP60783;  
 XX  
 XX DT 25-MAR-2003 (revised)  
 XX 23-OCT-1991 (first entry)  
 XX  
 XX DE Human urokinase.  
 XX  
 XX KW E.coli; high molecular urokinase.  
 XX  
 XX OS Homo sapiens.  
 XX  
 XX FH Key Location/Qualifiers  
 XX FT Protein 21..431  
 XX  
 XX DN JP61181377-A.  
 XX  
 XX PD 14-AUG-1986.  
 XX  
 XX PF 25-JAN-1985; 85JP-00011032.  
 XX  
 XX PR 25-JAN-1985; 85JP-00011032.  
 XX  
 XX PA (NISC ) NISSAN CHEM IND LTD.  
 XX (HODO ) HODOGAYA CHEM IND CO LTD.  
 XX (SAGA ) SAGAMI CHEM RES CENTRE.  
 XX (CENG ) CENTRAL GLASS CO LTD.  
 XX (NIPS ) NIPPON SODA CO.  
 XX (TOYU ) TOYO SODA MFG CO LTD.  
 XX  
 XX WPI; 1986-254744/39.  
 XX N-PSDB; AAN60703.

```

XX Human urokinase gene - has N-end of aminoacid sequence coded by codon
PT used in Escherichia coli.
PT
XX
XX Disclosure; Fig 2; 19pp; Japanese.
XX
XX The claimed gene product may be expressed in a transformed E.coli host,
CC for the efficient production of high molecular human urokinase. The N-
CC terminal of the protein expressed by the transforming plasmid is replaced
CC with a codon frequently used in E.coli. (Updated on 25-MAR-2003 to
CC correct PA field.)
XX
XX Sequence 431 AA;
SQ
Query Match 100.0%; Score 793; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 9.1e-54;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKFGGQHCETDKSKTCYEGNGHFYRG 60
DB 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKFGGQHCETDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRPNDRRPPWCYVQVGLK 120
DB 81 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRPNDRRPPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK 135
DB 141 PLVQECMVHDCADGK 155
RESULT 39
AAP70258
ID AAP70258 standard; protein; 431 AA.
XX
XX AAP70258;
XX
XX 25-MAR-2003 (revised)
DT 19-MAY-1991 (first entry)
XX
XX Sequence of human prourokinase and leader.
XX
XX Cardiovascular disease treatment; fibrin affinity; thrombolytic; enzyme;
XX protease.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Peptide 1..20
FT Protein /label= leader
FT Protein 21..431
FT Protein /label= prourokinase
XX
XX EP231883-A.
XX
XX 12-AUG-1987.
XX
XX 29-JAN-1987; 87EP-00101209.
XX
XX 31-JAN-1986; 86JP-00017734.
XX 30-JAN-1987; 87JP-00018626.
XX
XX (SAGA ) SAGAMI CHEM RES CENTRE.
XX (NIPS ) NIPPON SODA CO.
XX (CENG ) CENTRAL GLASS CO LTD.
XX (TOYJ ) TOYO SODA MFG CO LTD.
XX (NISC ) NISSAN CHEM IND LTD.
XX (NISC ) NISSAN CHEMICAL INDS KK.
XX
XX Tagawa M, Wada M, Yamada M, Yokoyama M, Numao N;
PI WPI; 1987-222882/32.
XX
XX N-ESDB; AAN70390.
XX

```

---

```

XX Hybrid plasminogen activator-like polypeptide - having a region for
PT affinity to fibrin from tissue plasminogen activator and a region from
PT prourokinase.
XX
XX Disclosure; Fig 2(1-5); 64pp; English.
XX
XX The TPA portion of the claimed hybrid polypeptide (see PT) may consist of
CC 2 kringles from N-terminal first serine to 219th glycine of human TPA, 1
CC kringles from 128th serine to 219th glycine of human TPA or half a kringles
CC from 161st methionine to 219th glycine (see AAP70257). The C-terminal
CC half of the hybrid polypeptide may contain an AA sequence from 150th
CC glutamine to C-terminal 41th leucine of prourokinase (see AAP70258).
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 431 AA;
SQ
Query Match 100.0%; Score 793; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 9.1e-54;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKFGGQHCETDKSKTCYEGNGHFYRG 60
DB 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKFGGQHCETDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRPNDRRPPWCYVQVGLK 120
DB 81 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRPNDRRPPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK 135
DB 141 PLVQECMVHDCADGK 155
RESULT 40
AAP71491
ID AAP71491 standard; protein; 431 AA.
XX
XX AAP71491;
XX
XX 31-OCT-2002 (revised)
DT 15-MAY-1991 (first entry)
XX
XX Modified prourokinase.
XX
XX Fibrin selectivity.
XX
XX Unidentified.
XX
XX Key Location/Qualifiers
FT Protein 21..431
FT Protein /label= Mature product
XX
XX EP236040-A.
XX
XX 09-SEP-1987.
XX
XX 24-FEB-1987; 87EP-00301567.
XX
XX 26-FEB-1986; 86US-00833179.
XX 19-FEB-1987; 87US-00012023.
XX
XX (COLB ) COLLABORATIVE RES INC.
XX
XX Vovis GF, Mao JI;
XX
XX WPI; 1987-251560/36.
XX
XX Amino acid modified form of prourokinase - having high fibrin selectively
PT and resistant to cleavage to the two-chain form.
XX
XX Claim 10; Fig 1; 18pp; Japanese.
XX

```

CC The modified prourokinase product has a single polypeptide chain, high  
CC fibrin selectivity and resistance to cleavage into the two-chain form.  
CC The modification is at the codon for Lys158. (Updated on 31-OCT-2002 to  
CC add missing OS field.)

XX SQ Sequence 431 AA;  
Query Match 100.0%; Score 793; DB 1; Length 431;  
Best Local Similarity 100.0%; Pred. No. 9.1e-54;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFRG 60  
DB 21 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFRG 80  
QY 61 KASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRRPPWCYVQGLK 120  
DB 81 KASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRRPPWCYVQGLK 140  
QY 121 PLVQECMVHDCADGK 135  
DB 141 PLVQECMVHDCADGK 155

RESULT 41  
AAP71663  
ID AAP71663 standard; protein; 431 AA.  
AC AAP71663;  
DT 30-APR-1991 (first entry)  
XX Modified prourokinase.  
KW Prourokinase; fibrin; PUK.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT Peptide 1..20 /label= sig\_peptide  
FT Protein 21..431 /label= mat\_protein  
FT Misc-difference 178..178 /note= "amino acid other than LYS"  
XX  
PN EP236040-A.  
XX  
XX 09-SEP-1987.  
XX  
XX 24-FEB-1987; 87EP-00301567.  
XX  
XX 26-FEB-1986; 86US-00833179.  
XX 19-FEB-1987; 87US-00012023.  
XX  
XX (COLB ) COLLABORATIVE RES INC.  
XX  
XX Vovis GF, Mao JI;  
XX  
XX WPI; 1987-251560/36.  
XX N-PSDB; AAN71338.  
XX  
XX Amino acid modified form of prourokinase - having high fibrin selectively  
XX and resistant to cleavage to the two-chain form.  
XX  
XX Disclosure; Fig 1; 18pp; Japanese.

XX The amino acid LYS at position 158 of the mature protein (see tag b) is  
XX replaced by another amino acid, achieved by site-directed mutagenesis or  
XX by chemical modification. The modified PUK has a single polypeptide chain,  
XX high fibrin activity and is resistant to cleavage to the two-chain form  
XX to an extent greater than its original unmodified form. The therapeutic  
XX specificity of single-chain PUK can be increased without decreasing

CC fibrinolytic activity. The single chain form can have high fibrin  
CC activity and thus does not cause general bleeding in the body but rather  
CC can be localised to the site of the blood clot directly. See also  
CC AAN71368-69

XX SQ Sequence 431 AA;  
Query Match 100.0%; Score 793; DB 1; Length 431;  
Best Local Similarity 100.0%; Pred. No. 9.1e-54;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFRG 60  
DB 21 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFRG 80  
QY 61 KASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRRPPWCYVQGLK 120  
DB 81 KASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRRPPWCYVQGLK 140  
QY 121 PLVQECMVHDCADGK 135  
DB 141 PLVQECMVHDCADGK 155

RESULT 42  
AAP71698  
ID AAP71698 standard; protein; 431 AA.  
XX  
XX AAP71698;  
XX  
XX 30-APR-1991 (first entry)  
XX Modified prourokinase (MET/SER).  
XX  
XX Prourokinase; fibrin; PUK.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FT Peptide 1..20 /label= sig\_peptide  
FT Protein 21..431 /label= mat\_protein  
XX  
XX EP236040-A.  
XX  
XX 09-SEP-1987.  
XX  
XX 24-FEB-1987; 87EP-00301567.  
XX  
XX 26-FEB-1986; 86US-00833179.  
XX 19-FEB-1987; 87US-00012023.  
XX  
XX (COLB ) COLLABORATIVE RES INC.  
XX  
XX Vovis GF, Mao JI;  
XX  
XX WPI; 1987-251560/36.  
XX N-PSDB; AAN71368.  
XX  
XX Amino acid modified form of prourokinase - having high fibrin selectively  
XX and resistant to cleavage to the two-chain form.  
XX  
XX Disclosure; Fig 1; 18pp; Japanese.

XX The amino acid LYS at position 158 of the mature protein is replaced by  
XX MET and the amino acid ILE at position 160 is replaced by SER, achieved  
XX by site-directed mutagenesis or by chemical modification. The modified PUK  
XX has a single polypeptide chain, high fibrin activity and is resistant to  
XX cleavage to the two-chain form to an extent greater than its original  
XX unmodified form. The therapeutic specificity of single-chain PUK can be  
XX increased without decreasing fibrinolytic activity. The single chain form  
XX can have high fibrin activity and thus does not cause general bleeding in



CC the body but rather can be localised to the site of the blood clot  
CC directly. See also AAN71368-69 and AAN71398  
XX  
SQ Sequence 431 AA;  
Query Match 100.0%; Score 793; DB 1; Length 431;  
Best Local Similarity 100.0%; Pred. No. 9.1e-54;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNELHGVPSNCDCLNGTGVSNKYFSNIHWNCNPKFGGQHCHEIDKSKTCYEGNGHFRG 60  
Db 21 SNELHGVPSNCDCLNGTGVSNKYFSNIHWNCNPKFGGQHCHEIDKSKTCYEGNGHFRG 80  
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120  
Db 81 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 140  
QY 121 PLVQECMVHDCADGK 135  
Db 141 PLVQECMVHDCADGK 155  
RESULT 43  
AAP71699  
ID AAP71699 standard; protein; 431 AA.  
AC AAP71699;  
XX  
DT 30-APR-1991 (first entry)  
DE Modified prourokinase (ALA).  
XX  
KW Prourokinase; fibrin; PUK.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..20  
FT Protein /label= sig\_peptide 21..431  
FT /label= mat\_protein  
XX  
PN EF236040-A.  
XX  
PD 09-SEP-1987.  
XX  
PF 24-FEB-1987; 87EP-00301567.  
XX  
PR 26-FEB-1986; 86US-00833179.  
PR 19-FEB-1987; 87US-00012023.  
XX  
PA (COLB ) COLLABORATIVE RES INC.  
XX  
PI Vovis GF, Mao JI;  
XX  
DR WPI; 1987-251560/36.  
DR N-PSDB; AAN71369.  
XX  
PT Amino acid modified form of prourokinase - having high fibrin selectively  
XX and resistant to cleavage to the two-chain form.  
PS Disclosure; Fig 1; 18pp; Japanese.  
XX  
CC The amino acid Lys at position 158 of the mature protein (bases 610-612)  
CC is replaced by another amino acid (ALA), achieved by site-directed  
CC mutagenesis. The modified PUK has a single polypeptide chain, high fibrin  
CC activity and is resistant to cleavage to the two-chain form to an extent  
CC greater than its original unmodified form. The therapeutic specificity of  
CC single-chain PUK can be increased without decreasing fibrinolytic  
CC activity. The single chain form can have high fibrin activity and thus  
CC does not cause general bleeding in the body but rather can be localised  
CC to the site of the blood clot directly. See also AAN71368-69  
XX

SQ Sequence 431 AA;  
Query Match 100.0%; Score 793; DB 1; Length 431;  
Best Local Similarity 100.0%; Pred. No. 9.1e-54;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNELHGVPSNCDCLNGTGVSNKYFSNIHWNCNPKFGGQHCHEIDKSKTCYEGNGHFRG 60  
Db 21 SNELHGVPSNCDCLNGTGVSNKYFSNIHWNCNPKFGGQHCHEIDKSKTCYEGNGHFRG 80  
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120  
Db 81 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 140  
QY 121 PLVQECMVHDCADGK 135  
Db 141 PLVQECMVHDCADGK 155  
RESULT 44  
AAP80430  
ID AAP80430 standard; protein; 431 AA.  
XX  
AC AAP80430;  
XX  
DT 25-MAR-2003 (revised)  
DT 14-SEP-1990 (first entry)  
XX  
DE Deduced AA sequence of the single chain urokinase plasminogen activator  
XX (SCU-PA) cDNA insert prepared from human Hep3 cells.  
XX  
KW Single chain urokinase plasminogen activator (SCU-PA); human Hep3 cells;  
KW glyceraldehyde-3-phosphate dehydrogenase gene; thrombosis prevention;  
KW thrombosis treatment.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..20  
FT Protein 21..411  
XX  
PN EP288435-A.  
XX  
PD 26-OCT-1988.  
XX  
PF 11-APR-1988; 88EP-00810234.  
XX  
PR 15-APR-1987; 87GB-00009081.  
PR 16-JUN-1987; 87GB-00014059.  
PR 04-DEC-1987; 87TE-00003299.  
XX  
PA (CIBA ) CIBA GEIGY AG.  
XX  
PI Meyhack B, Heim J, Burgi R;  
XX  
DR WPI; 1988-301440/43.  
DR N-PSDB; AAN80981.  
XX  
PT Prodn. of human single chain urokinase-type plasminogen activator - by  
XX culturing yeast strain transformed with hybrid vector con9. Yeast  
XX expression control sequences.  
PS Example 1; Fig 2; 48pp; English.  
XX  
CC The patent is for the prodn. of human single chain urokinase-type  
CC plasminogen activator (UTPA). Mutants of scu-PA are especially those  
CC which render the protein protease resistant. Such scu-PA mutants are  
CC covalently modified at sites of proteolysis by proteases occurring in  
CC blood such as thrombin or plasmin, so that they are no longer susceptible  
CC to protease hydrolysis at these locations. The target sites include  
CC Lys135 to Lys136 (cleavage at this site generates the so-called low  
CC molecular weight form of scu-PA or LUX), Arg156 to Phe157 (susceptible to  
CC thrombin attack) and Lys158 to Ile159 (cleavage at this site by plasmin

CC Generates tuc-PA). Suitable scu-PA mutants have site specific  
 CC substitutions, insertions or deletions of residues at one or more of  
 CC these target sites. Especially preferred are those mutants in which one  
 CC amino acid residue or both amino acid residues forming the target sites  
 CC are deleted or in which at least one of these amino acid residues is  
 CC replaced by another amino acid residue so that the resulting mutants are  
 CC resistant to proteolytic attack. The UPA proteins exhibit the biological  
 CC activity of natural human UPA without any refolding procedure being  
 CC necessary. They can be used as for known PAs in humans for the prevention  
 CC or treatment of thrombosis or other conditions where it is desired to  
 CC produce local fibrinolytic or proteolytic activity. (Updated on 25-MAR-  
 CC 2003 to correct PR field.)  
 XX SQ Sequence 431 AA;

Query Match 100.0%; Score 793; DB 1; Length 431;  
 Best Local Similarity 100.0%; Pred. No. 9.1e-54;  
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKXFGGQHCIDKSKTCYEGNGHFYRG 60  
 Db 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKXFGGQHCIDKSKTCYEGNGHFYRG 80  
 QY 61 KASDTMTGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRPNRRRPWCYVQVGLK 120  
 Db 81 KASDTMTGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRPNRRRPWCYVQVGLK 140  
 QY 121 PLVOECMVHDCADGK 135  
 Db 141 PLVOECMVHDCADGK 155

RESULT 45  
 AAP81204  
 ID AAP81204 standard; protein; 431 AA.  
 AC AAP81204;  
 XX 25-MAR-2003 (revised)  
 DT 03-DEC-1990 (first entry)  
 XX Pro-urokinase with signal sequence.  
 XX pro-urokinase (pro-UK); plasminogen activator; pUK33; ss.  
 XX Homo sapiens.  
 XX Key Location/Qualifiers  
 XX Peptide 1..20  
 FT /label= signal peptide  
 FT Protein 21..431  
 FT /label= pro-urokinase  
 XX EP265874-A.  
 XX 04-MAY-1988.  
 XX 23-OCT-1987; 87EP-00115600.  
 XX 23-OCT-1986; 86JP-00253078.  
 XX (GREC ) GREEN CROSS CORP.  
 XX Anatsuji Y, Okabayashi K, Nagai M, Arimura H, Suyama T;  
 XX WPI; 1988-121000/18.  
 XX N-PSDB; AAN81558.  
 XX Glycosylated single-chain pro-urokinase prodn. - by cultivating DHFR gene  
 XX -deficient CHO-K1 cells transformed with a plasmid contg. CDNA, SV40  
 XX promoter and DHFR gene.  
 XX Disclosure; Page ?; 19pp; English.

XX The Arg at position 2 is encoded by TGA(sic). Possible error in the  
 CC specification. Should read CGA? The pro-UK gene was derived from plasmid  
 CC pUK33. The cDNA was synthesised using urokinase mRNA isolated from a  
 CC human kidney cell line. Pro-UK was cloned into a SV40 promoter-contg.  
 CC plasmid, down- stream of the promoter. This plasmid was then ligated to  
 CC a DHFR- gene contg. plasmid so that pro-UK and DHFR are inserted in  
 CC opposite directions. The recombinant plasmid was used to transform CHO-K1  
 CC cell derived DHFR gene-deficient host cells to produce glycosylated  
 CC single-chain pro-UK. (Updated on 25-MAR-2003 to correct PF field.)  
 XX SQ Sequence 431 AA;

Query Match 100.0%; Score 793; DB 1; Length 431;  
 Best Local Similarity 100.0%; Pred. No. 9.1e-54;  
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKXFGGQHCIDKSKTCYEGNGHFYRG 60  
 Db 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKXFGGQHCIDKSKTCYEGNGHFYRG 80  
 QY 61 KASDTMTGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRPNRRRPWCYVQVGLK 120  
 Db 81 KASDTMTGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRPNRRRPWCYVQVGLK 140  
 QY 121 PLVOECMVHDCADGK 135  
 Db 141 PLVOECMVHDCADGK 155

RESULT 46  
 AAP91886  
 ID AAP91886 standard; protein; 431 AA.  
 AC AAP91886;  
 XX 25-MAR-2003 (revised)  
 DT 31-OCT-2002 (revised)  
 DT 16-APR-1990 (first entry)  
 XX Sequence of prourokinase.  
 XX Low mol. wt. plasminogen activator; prourokinase; dissolve blood clot.  
 XX Homo sapiens.  
 XX Key Location/Qualifiers  
 XX Protein 21..431  
 FT /note= "Mature prourokinase."  
 FT Region 170..179  
 FT /note= "Preferred initiation region for the low mol. wt.  
 FT plasminogen activators."  
 XX EP316068-A.  
 XX 17-MAY-1989.  
 XX 07-OCT-1988; 88EP-00309417.  
 XX 09-OCT-1987; 87US-00107370.  
 XX 27-SEP-1988; 88US-00248727.  
 XX (COLB ) COLLABORATIVE RES INC.  
 XX Mao JI;  
 XX WPI; 1989-146601/20.  
 XX N-PSDB; AAN91740.  
 XX Modified low mol. wt. plasminogen activator - formed of amino acids  
 XX comprising the amino acid portion of prourokinase from 150 to 411.  
 XX Fig 1; Page -; 27pp; English.

XX A low mol. wt. polypeptide plasminogen activator (PA) is claimed which is  
CC formed from amino acids 150-411 of prourokinase. The preferred initiation  
CC region for the low mol. wt. PA is indicated (see FT). A low mol wt. PA  
CC can be injected into blood in the body in vivo to dissolve clots without  
CC harm. (Updated on 31-OCT-2002 to add missing OS field.) (Updated on 25-  
CC MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PI  
CC field.)  
XX  
XX Sequence 431 AA;  
SQ  
Query Match 100.0%; Score 793; DB 1; Length 431;  
Best Local Similarity 100.0%; Pred. No. 9.1e-54;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNEHQVPSNCDCLNGGTCVSNKYPFNIHWNCNPKFGGQHCETDKSKTCYEGNGHFYRG 60  
Db 21 SNEHQVPSNCDCLNGGTCVSNKYPFNIHWNCNPKFGGQHCETDKSKTCYEGNGHFYRG 80  
QY 61 KASDTMTGRPCLPNNSATVLQOTYHAHRSALQGLGKHNYCRNPDNRPRPWCYVQVGLK 120  
Db 81 KASDTMTGRPCLPNNSATVLQOTYHAHRSALQGLGKHNYCRNPDNRPRPWCYVQVGLK 140  
QY 121 PLVQECMVHDCADGK 135  
Db 141 PLVQECMVHDCADGK 155  
RESULT 47  
AAP92119  
ID AAP92119 standard; protein; 431 AA.  
XX  
AC AAP92119;  
XX  
XX 25-MAR-2003 (revised)  
DT 29-JUN-1990 (first entry)  
XX  
DE Natural human prourokinase.  
XX  
KW Human prourokinase; antithrombotic; derivative.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FT Misc-difference 1 /note= "Optional in new deriv."  
FT Misc-difference 2..155  
FT Misc-difference /note= "Incorporated into new deriv."  
FT Misc-difference 135  
FT Misc-difference /note= "May be replaced by a non-basic AA in new deriv."  
FT Misc-difference 156  
FT Misc-difference /note= "Undefined residue in new deriv."  
FT Misc-difference 157  
FT Misc-difference /note= "Pro, Gly, Ala or Val in new deriv."  
FT Misc-difference 158  
FT Misc-difference /note= "Lys or Arg in new deriv."  
XX  
XX WO8901513-A.  
XX  
XX 23-FEB-1989.  
XX  
XX 18-AUG-1988; 89WO-JP000815.  
XX  
XX 19-AUG-1987; 87JP-00204149.  
XX  
XX (SAGA ) SAGAMI CHEM RES CENTRE.  
PA (CENG ) CENTRAL GLASS CO LTD.  
PA (HODO ) HODOGAYA CHEM KK.  
PA (NIPS ) NIPPON SODA CO.  
PA (NISC ) NISSAN CHEM IND LTD.  
XX Kobayashi Y, Omori M, Yamada C;  
PI  
XX

DR WPI; 1989-068869/09.  
DR N-PSDB; AAN91075.  
XX  
XX Antithrombotic fast-acting pro-urokinase deriv. - produced by culture of  
PT E coli transformant contg. new plasmid of PMUT9Q family.  
XX  
XX Disclosure; Fig 1; 75pp; Japanese.  
XX  
XX A human prourokinase (PU) deriv. is new which is based upon residues 2-  
CC 155 of natural human prourokinase. The new deriv. is produced by E. coli  
CC J103/PMUT90-RPK in culture. It is a fast-acting drug for the treatment  
CC and prevention of thrombosis. (Updated on 25-MAR-2003 to correct PA  
CC field.) (Updated on 25-MAR-2003 to correct PI field.)  
XX  
XX Sequence 431 AA;  
SQ  
Query Match 100.0%; Score 793; DB 1; Length 431;  
Best Local Similarity 100.0%; Pred. No. 9.1e-54;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0  
QY 1 SNEHQVPSNCDCLNGGTCVSNKYPFNIHWNCNPKFGGQHCETDKSKTCYEGNGHFYRG 60  
Db 21 SNEHQVPSNCDCLNGGTCVSNKYPFNIHWNCNPKFGGQHCETDKSKTCYEGNGHFYRG 80  
QY 61 KASDTMTGRPCLPNNSATVLQOTYHAHRSALQGLGKHNYCRNPDNRPRPWCYVQVGLK 120  
Db 81 KASDTMTGRPCLPNNSATVLQOTYHAHRSALQGLGKHNYCRNPDNRPRPWCYVQVGLK 140  
QY 121 PLVQECMVHDCADGK 135  
Db 141 PLVQECMVHDCADGK 155  
RESULT 48  
AAP94764  
ID AAP94764 standard; protein; 431 AA.  
XX  
AC AAP94764;  
XX  
XX 25-MAR-2003 (revised)  
DT 27-JUN-1990 (first entry)  
XX  
XX Non-glycosylated prourokinase.  
XX  
XX Prourokinase; CGE 195; plasminogen activator; blood clot lysis.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FT Region 1..20  
FT /label= signal sequence  
XX  
XX EP299706-A.  
XX  
XX 18-JAN-1989.  
XX  
XX 12-JUL-1988; 88EP-00306334.  
XX  
XX 13-JUL-1987; 87US-00072426.  
XX 29-JUN-1988; 88US-00211279.  
XX  
XX (COLB ) COLLABORATIVE RES INC.  
XX  
XX Baltimore D, Moir DT, Broeze RJ;  
XX  
XX WPI; 1989-017204/03.  
DR N-PSDB; AAN93079.  
XX  
XX New non glycosylated, secreted plasminogen activator - pref. with  
PT asparagine replaced or deleted, useful for treating blood clots.  
PT expressed in non mammalian cells.  
XX  
XX Disclosure; Page; 26pp; English.  
PS

```

XX myocardial infarction. DNA encoding the protein was sequenced from
CC plasmid pGGE195, a subclone of two inserts isolated by screening a cDNA
CC library prep'd. from kidney cell RNA. One of the original inserts, clone
CC CGF31 (tag c) started in the middle of the signal sequence. Mutants of
CC the sequence, pref. in which gcc (Ala) replaces aat (Asn) at nucleotides
CC 1002-1004 (residue 302) are used to transform hosts for the prodn. of non
CC -glycosylated pro-urokinase. (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 431 AA;
Query Match 100.0%; Score 793; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 9.1e-54;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNCPKFGQHCIEDKSKTCYEGNGHFYRG 60
Db 21 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNCPKFGQHCIEDKSKTCYEGNGHFYRG 80
QY 61 KASDTDMGRPCLPWNSATVLQOQTYHAHRS DALQLGLGKHNYCRNPDNRRPWCYVQVGLK 120
Db 81 KASDTDMGRPCLPWNSATVLQOQTYHAHRS DALQLGLGKHNYCRNPDNRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK 135
Db 141 PLVQECMVHDCADGK 155
RESULT 49
AAR07112
ID AAR07112 standard; protein; 431 AA.
XX
AC AAR07112;
XX
DT 25-MAR-2003 (revised)
DT 24-JAN-1991 (first entry)
XX
DE Human pro-Urokinase encoded by plasmid pUK1.
XX
KW pro-Urokinase; transgenic mice.
XX
OS Synthetic.
XX
FN EP390592-A.
XX
PD 03-OCT-1990.
XX
PF 30-MAR-1990; 90EP-00303445.
XX
PR 31-MAR-1989; 89JP-00078574.
XX
PA (KYOW ) KYOWA HAKKO KOGYO KK.
PA (EXPE-) CENT INST EXPR ANIMALS.
PA (JIKK-) JIKKEN DOBUTSU CHUO KENK.
XX
PI Sekine S, Ito S, Katsuki M;
XX
WPI; 1990-299492/40.
DR N-PSDB; AAQ06049.
XX
PT Prodn. of recombinant protein, esp. human pro-urokinase - from milk of
PT transgenic animals using promoter of bovine alpha S1 casein chromosomal
PT gene.
XX
PS Example; Table 1; 55pp; English.
XX
CC E.coli strain C600SF8 was transformed with recombinant plasmid containing
CC ds-DNA derived from human pharynx cancer cell strain Detroit 562. 10000
CC colonies were screened and one positive clone was identified. Plasmid
CC pUK1 was isolated and found to contain the coding region and 3' non-
CC coding region of pro-UK downstream of Cys(41). Four silent substitutions
CC were identified c.f. Holmes et al., Biotechnology, vol 3, p.923 (1985) as
CC follows: (sic) Asn(254), AAC to AAT; Leu(340), CTA to CTG; Pro(345), CCC

```

```

CC to CCA; Gln(346), CAA to CAG. See also AAQ06045-Q06048 and AAQ06392.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 431 AA;
Query Match 100.0%; Score 793; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 9.1e-54;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNCPKFGQHCIEDKSKTCYEGNGHFYRG 60
Db 21 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNCPKFGQHCIEDKSKTCYEGNGHFYRG 80
QY 61 KASDTDMGRPCLPWNSATVLQOQTYHAHRS DALQLGLGKHNYCRNPDNRRPWCYVQVGLK 120
Db 81 KASDTDMGRPCLPWNSATVLQOQTYHAHRS DALQLGLGKHNYCRNPDNRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK 135
Db 141 PLVQECMVHDCADGK 155
RESULT 50
AAR04253
ID AAR04253 standard; protein; 431 AA.
XX
AC AAR04253;
XX
DT 25-MAR-2003 (revised)
DT 12-SEP-1990 (first entry)
XX
DE Human pro-urokinase from the cDNA of clone pcUK176.
XX
KW Non-glycosylated; pro-urokinase; E. coli; Ptcp promoter; MS-2 RBS.
XX
OS Synthetic.
XX
FN EP365894-A.
XX
PD 02-MAY-1990.
XX
PF 06-OCT-1989; 89EP-00118586.
PR 11-OCT-1988; 88GB-00023833.
XX
PA (FARM ) FARMITALIA ERBA SPA CARLO.
XX
PI Brandazza A, Sarmientos P, Orsini G;
XX
WPI; 1990-133447/18.
DR N-PSDB; AAQ04107.
XX
PT Non-glycosylated pro-urokinase prodn. - using E coli B strains and E coli
PT promoter PTRP and Shine-Dalgarno sequence MS-2.
XX
PS Disclosure; Page ?; -pp; English.
XX
CC SER residue at position 21 is the start of the mature proUK. Non-
CC glycosylated proUK (MW 45kd) produced by E.coli B strain containing the
CC sequence. See also AAQ04101-07. (Updated on 25-MAR-2003 to correct PA
CC field.)
XX
SQ Sequence 431 AA;
Query Match 100.0%; Score 793; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 9.1e-54;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNCPKFGQHCIEDKSKTCYEGNGHFYRG 60
Db 21 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNCPKFGQHCIEDKSKTCYEGNGHFYRG 80
QY 61 KASDTDMGRPCLPWNSATVLQOQTYHAHRS DALQLGLGKHNYCRNPDNRRPWCYVQVGLK 120

```

Db 81 KASTDTMGRPCLPWNSAIVLQOTYHAHRS DALQLGLGKNTCRNPDNRPRPCYVQGLK 140  
Oy 121 PLVQECMVHDCADGK 135  
Db 141 PLVQECMVHDCADGK 155

Search completed: May 25, 2004, 14:52:54  
Job time : 36.0105 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2004, 14:53:05 ; Search time 24.61 Seconds  
(without alignments)  
1530.046 Million cell updates/sec

Title: US-09-880-503-4  
Perfect score: 793  
Sequence: 1 SNELHQVPSNCDCLNGTCV.....QVGLKPLVQECVHDCADGK 135

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1149313 seqs, 278921704 residues

Total number of hits satisfying chosen parameters: 1149313

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pap.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pap.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pap.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pap.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pap.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pap.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pap.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pap.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pap.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pap.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pap.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pap.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pap.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pap.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pap.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pap.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pap.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	793	100.0	135	9	US-09-880-503-4
2	793	100.0	138	9	US-09-984-186-12
3	793	100.0	138	14	US-10-237-667-12
4	793	100.0	138	14	US-10-237-708-12
5	793	100.0	138	14	US-10-237-866-12
6	793	100.0	138	14	US-10-237-871-12
7	793	100.0	138	14	US-10-237-871-12
8	793	100.0	138	14	US-10-237-871-12
9	793	100.0	138	16	US-10-702-536-12
10	793	100.0	138	16	US-10-702-536-12
11	793	100.0	137	9	US-09-880-503-8
12	793	100.0	337	14	US-10-106-698-6266
13	793	100.0	337	15	US-10-264-049-2927
14	793	100.0	403	9	US-09-880-503-6
15	793	100.0	411	9	US-09-880-503-3
16	793	100.0	431	9	US-09-264-468B-1

## RESULT 1

US-09-880-503-4  
; Sequence 4, Application US/09880503  
; Patent No. US20020131964A1  
; GENERAL INFORMATION:  
; APPLICANT: CINES, Douglas B  
; APPLICANT: HIGAZI, Abd Al-Roof  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL ANI

## ALIGNMENTS

16	793	100.0	431	12	US-10-411-037-34	Sequence 34, Appl
17	793	100.0	431	12	US-10-411-026-34	Sequence 34, Appl
18	793	100.0	431	13	US-10-076-421-2	Sequence 2, Appl
19	793	100.0	431	14	US-10-171-311-184	Sequence 184, App
20	793	100.0	431	14	US-10-193-656-4	Sequence 4, Appl
21	793	100.0	431	14	US-10-301-822-161	Sequence 161, App
22	793	100.0	431	14	US-10-131-985-21	Sequence 21, Appl
23	793	100.0	431	15	US-10-295-027-414	Sequence 414, App
24	793	100.0	431	15	US-10-235-027-1275	Sequence 1275, A
25	793	100.0	431	16	US-10-410-362-34	Sequence 34, Appl
26	793	100.0	431	16	US-10-411-049-34	Sequence 34, Appl
27	793	100.0	437	12	US-10-087-192-594	Sequence 594, App
28	791	99.7	431	14	US-10-247-671-149	Sequence 149, App
29	783	98.7	411	15	US-10-407-821-2	Sequence 2, Appl
30	777	98.0	431	12	US-10-282-174-562	Sequence 562, App
31	771	97.2	445	15	US-10-360-101-266	Sequence 266, App
32	595.5	75.1	650	15	US-10-401-077-1	Sequence 1, Appl
33	548	69.1	433	12	US-10-087-192-591	Sequence 591, App
34	510	64.3	88	9	US-09-880-503-1	Sequence 1, Appl
35	510	64.3	96	9	US-09-880-503-9	Sequence 9, Appl
36	501	63.2	91	12	US-10-424-999-37	Sequence 37, Appl
37	501	63.2	221	12	US-10-424-999-36	Sequence 36, Appl
38	501	63.2	687	12	US-10-424-999-17	Sequence 17, Appl
39	501	63.2	687	12	US-10-425-000-37	Sequence 37, Appl
40	501	63.2	687	15	US-10-233-675A-17	Sequence 17, Appl
41	500	63.1	87	12	US-10-424-999-9	Sequence 9, Appl
42	500	63.1	87	12	US-10-425-000-29	Sequence 29, Appl
43	500	63.1	87	15	US-10-233-675A-9	Sequence 9, Appl
44	499	62.9	86	12	US-10-424-999-5	Sequence 5, Appl
45	499	62.9	86	12	US-10-424-999-62	Sequence 62, Appl
46	499	62.9	86	12	US-10-425-000-97	Sequence 97, Appl
47	499	62.9	86	15	US-10-233-675A-5	Sequence 5, Appl
48	499	62.9	86	15	US-10-233-675A-22	Sequence 22, Appl
49	498	62.8	688	12	US-10-424-999-18	Sequence 18, Appl
50	498	62.8	688	12	US-10-425-000-38	Sequence 38, Appl
51	498	62.8	688	15	US-10-233-675A-18	Sequence 18, Appl
52	498	62.8	689	12	US-10-424-999-13	Sequence 13, Appl
53	498	62.8	689	12	US-10-425-000-32	Sequence 32, Appl
54	498	62.8	689	15	US-10-233-675A-13	Sequence 13, Appl
55	496	62.5	91	12	US-10-424-999-35	Sequence 35, Appl
56	496	62.5	221	12	US-10-424-999-34	Sequence 34, Appl
57	495	62.4	87	12	US-10-424-999-10	Sequence 10, Appl
58	495	62.4	87	15	US-10-425-000-30	Sequence 30, Appl
59	495	62.4	87	15	US-10-233-675A-10	Sequence 10, Appl
60	495	62.4	322	12	US-10-424-999-21	Sequence 21, Appl
61	495	62.4	322	12	US-10-425-000-41	Sequence 41, Appl
62	495	62.4	322	15	US-10-233-675A-21	Sequence 21, Appl
63	495	62.4	672	12	US-10-424-999-15	Sequence 15, Appl
64	495	62.4	672	12	US-10-425-000-35	Sequence 35, Appl
65	495	62.4	672	15	US-10-233-675A-15	Sequence 15, Appl
66	495	62.4	674	12	US-10-424-999-14	Sequence 14, Appl
67	495	62.4	674	12	US-10-425-000-34	Sequence 34, Appl
68	495	62.4	674	15	US-10-233-675A-14	Sequence 14, Appl
69	494	62.3	86	12	US-10-424-999-1	Sequence 1, Appl
70	494	62.3	86	15	US-10-233-675A-1	Sequence 1, Appl
71	494	62.3	322	12	US-10-424-999-20	Sequence 20, Appl
72	494	62.3	322	12	US-10-425-000-40	Sequence 40, Appl
73	494	62.3	322	15	US-10-233-675A-20	Sequence 20, Appl
74	489	61.7	86	12	US-10-424-999-7	Sequence 7, Appl
75	489	61.7	86	15	US-10-233-675A-7	Sequence 7, Appl

;; TITLE OF INVENTION: TISSUE CONTRACTABILITY  
;; FILE REFERENCE: 9596-331  
;; CURRENT APPLICATION NUMBER: US/09/880,503  
;; CURRENT FILING DATE: 2001-05-13  
;; PRIOR APPLICATION NUMBER: US 60/212,847  
;; PRIOR FILING DATE: 2000-06-20  
;; NUMBER OF SEQ ID NOS: 18  
;; SOFTWARE: Patent In Ver. 2.1  
;; SEQ ID NO 4  
;; LENGTH: 135  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-880-503-4

Query Match 100.0%; Score 793; DB 9; Length 135;  
Best Local Similarity 100.0%; Pred. No. 3.6e-72;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60  
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60  
QY 61 KASTDTMGRPCLPWSATVLTQTYHAHRSALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120  
Db 61 KASTDTMGRPCLPWSATVLTQTYHAHRSALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120  
QY 121 PLVQECMVHDCADGK 135  
Db 121 PLVQECMVHDCADGK 135

## RESULT 2

US-09-984-186-12  
; Sequence 12, Application US/09984186  
; Patent No. US2002015101A1  
; GENERAL INFORMATION:

APPLICANT: Fleer, Reinhard  
Fournier, Alain  
Guitton, Jean-Dominique  
Jung, Gerard  
Yeh, Patrice

TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: System 7.1  
SOFTWARE: Word 5.1 (PatentIn)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/984,186  
FILING DATE: 29-Oct-2001  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/797,689  
FILING DATE: 31-JAN-1997  
APPLICATION NUMBER: US 08/256,927  
FILING DATE: 28-JUL-1994  
APPLICATION NUMBER: FR 92/01064  
FILING DATE: 31-JAN-1992  
APPLICATION NUMBER: PCT/FR93/00085  
FILING DATE: 28-JAN-1993

ATTORNEY/AGENT INFORMATION:  
NAME: Smith Ph.D., Julie K.  
REGISTRATION NUMBER: P-38,619

;; REFERENCE/DOCKET NUMBER: ST92006-US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (610) 454-3839  
;; TELEFAX: (610) 454-3808  
;; INFORMATION FOR SEQ ID NO: 12:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 138 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-09-984-186-12

Query Match 100.0%; Score 793; DB 9; Length 138;  
Best Local Similarity 100.0%; Pred. No. 3.7e-72;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60  
Db 4 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 63  
QY 61 KASTDTMGRPCLPWSATVLTQTYHAHRSALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120  
Db 64 KASTDTMGRPCLPWSATVLTQTYHAHRSALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 123  
QY 121 PLVQECMVHDCADGK 135  
Db 124 PLVQECMVHDCADGK 138

## RESULT 3

US-10-237-667-12  
; Sequence 12, Application US/10237667  
; Publication No. US20030022308A1  
; GENERAL INFORMATION:

APPLICANT: Fleer, Reinhard  
Fournier, Alain  
Guitton, Jean-Dominique  
Jung, Gerard  
Yeh, Patrice

TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA  
ZIP: 19426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: System 7.1  
SOFTWARE: Word 5.1 (PatentIn)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/237,667  
FILING DATE: 10-Sep-2002  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/797,689  
FILING DATE: 31-JAN-1997  
APPLICATION NUMBER: US 08/256,927  
FILING DATE: 28-JUL-1994  
APPLICATION NUMBER: FR 92/01064  
FILING DATE: 31-JAN-1992  
APPLICATION NUMBER: PCT/FR93/00085  
FILING DATE: 28-JAN-1993

ATTORNEY/AGENT INFORMATION:  
NAME: Smith Ph.D., Julie K.  
REGISTRATION NUMBER: P-38,619  
REFERENCE/DOCKET NUMBER: ST92006-US

Query Match	100.0%	Score 793	DB 14	Length 138
Best Local Similarity	100.0%	Pred. No. 3.7e-72		
Matches 135	Conservative 0	Mismatches 0	Indels 0	Gaps 0
US-10-237-708-12				
QY	1	SNELHVPNCCLNGGTCVSNKYFSNIHWCNPKKFGQHQHCEIDKSKTCYEGNGHFFYRG	60	
Db	4	SNELHVPNCCLNGGTCVSNKYFSNIHWCNPKKFGQHQHCEIDKSKTCYEGNGHFFYRG	63	
QY	61	KASTDTMGPRCLPWSNATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRPRRRCYVQVGLK	120	
Db	64	KASTDTMGPRCLPWSNATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRPRRRCYVQVGLK	123	
QY	121	PLVQECMVHDCADGK	135	
Db	124	PLVQECMVHDCADGK	138	

RESULT 4

US-10-237-708-12

Sequence 12, Application US/10237708

Publication No. US20030036170A1

GENERAL INFORMATION:

APPLICANT: Fleer, Reinhard

Fournier, Alain

Guittion, Jean-Dominique

Jung, Gerard

Yeh, Patrice

TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES, PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESSEE: Rhone-Poulenc Rorer Inc.

STREET: 500 Arcola Road, 3C43

CITY: Collegeville

STATE: PA

COUNTRY: USA

ZIP: 19426

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: Macintosh

OPERATING SYSTEM: System 7.1

SOFTWARE: Word 5.1 (PatentIn)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/237,708

FILING DATE: 10-Sep-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/797,689

FILING DATE: 31-JAN-1997

APPLICATION NUMBER: US 08/256,927

FILING DATE: 28-JUL-1994

APPLICATION NUMBER: FR 92/01064

FILING DATE: 31-JAN-1992

APPLICATION NUMBER: PCT/FR93/00085

FILING DATE: 28-JAN-1993

ATTORNEY/AGENT INFORMATION:

NAME: Smith Ph.D., Julie K.

REGISTRATION NUMBER: P-38,619

REFERENCE/DOCKET NUMBER: ST92006-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (610) 454-3839

TELEFAX: (610) 454-3808

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 138 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-10-237-667-12

Query Match

Best Local Similarity

Matches 135

Conservative 0

Mismatches 0

Indels 0

Gaps 0

US-10-237-667-12

QY

1

SNELHVPNCCLNGGTCVSNKYFSNIHWCNPKKFGQHQHCEIDKSKTCYEGNGHFFYRG

60

Db

4

SNELHVPNCCLNGGTCVSNKYFSNIHWCNPKKFGQHQHCEIDKSKTCYEGNGHFFYRG

63

QY

61

KASTDTMGPRCLPWSNATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRPRRRCYVQVGLK

120

Db

64

KASTDTMGPRCLPWSNATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRPRRRCYVQVGLK

123

QY

121

PLVQECMVHDCADGK

135

Db

124

PLVQECMVHDCADGK

138



```
/
/ TELEFAX: (610) 454-3808
/ INFORMATION FOR SEQ ID NO: 12:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 138 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-866-12

Query Match      100.0%; Score 793; DB 14; Length 138;
Best Local Similarity 100.0%; Pred. No. 3,7e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKVFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFPYRG 60
DB 4 SNELHQVPSNCDCLNGGTCVSNKVFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFPYRG 63

QY 61 KASDTDMGRPCLPWSATVLQOTYHAHRSALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120
DB 64 KASDTDMGRPCLPWSATVLQOTYHAHRSALQLGLGKHNYCRPNDRRRPWCYVQVGLK 123

QY 121 PLVQECMVHDCADGK 135
DB 124 PLVQECMVHDCADGK 138

RESULT 6
US-10-237-871-12
/ Sequence 12, Application US/10237871
/ Publication No. US20030036172A1
/ GENERAL INFORMATION:
/ APPLICANT: Fleer, Reinhard
/ Fournier, Alain
/ Guittou, Jean-Dominique
/ Jung, Gerard
/ Yeh, Patrice
/ TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
/ PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
/ CONTAINING SAID POLYPEPTIDES
/ NUMBER OF SEQUENCES: 36
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Rhone-Poulenc Rorer Inc.
/ STREET: 500 Arcola Road, 3043
/ CITY: Collegeville
/ STATE: PA
/ COUNTRY: USA
/ ZIP: 19426
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: Macintosh
/ OPERATING SYSTEM: System 7.1
/ SOFTWARE: Word 5.1 (Patentin)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/237,871
/ FILING DATE: 10-Sep-2002
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/797,689
/ FILING DATE: 31-JAN-1997
/ APPLICATION NUMBER: US 08/256,927
/ FILING DATE: 28-JUL-1994
/ APPLICATION NUMBER: FR 92/01064
/ FILING DATE: 31-JAN-1992
/ APPLICATION NUMBER: PCT/FR93/00085
/ FILING DATE: 28-JAN-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Smith Ph.D., Julie K.
/ REGISTRATION NUMBER: P-38,619
/ REFERENCE/DOCKET NUMBER: ST92006-US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (610) 454-3839
/ TELEFAX: (610) 454-3808
```

```
/
/ INFORMATION FOR SEQ ID NO: 12:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 138 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-871-12

Query Match      100.0%; Score 793; DB 14; Length 138;
Best Local Similarity 100.0%; Pred. No. 3,7e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKVFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFPYRG 60
DB 4 SNELHQVPSNCDCLNGGTCVSNKVFNSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFPYRG 63

QY 61 KASDTDMGRPCLPWSATVLQOTYHAHRSALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120
DB 64 KASDTDMGRPCLPWSATVLQOTYHAHRSALQLGLGKHNYCRPNDRRRPWCYVQVGLK 123

QY 121 PLVQECMVHDCADGK 135
DB 124 PLVQECMVHDCADGK 138

RESULT 7
US-10-237-624-12
/ Sequence 12, Application US/10237624
/ Publication No. US20030082747A1
/ GENERAL INFORMATION:
/ APPLICANT: Fleer, Reinhard
/ Fournier, Alain
/ Guittou, Jean-Dominique
/ Jung, Gerard
/ Yeh, Patrice
/ TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
/ PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITIO
/ CONTAINING SAID POLYPEPTIDES
/ NUMBER OF SEQUENCES: 36
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Rhone-Poulenc Rorer Inc.
/ STREET: 500 Arcola Road, 3043
/ CITY: Collegeville
/ STATE: PA
/ COUNTRY: USA
/ ZIP: 19426
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: Macintosh
/ OPERATING SYSTEM: System 7.1
/ SOFTWARE: Word 5.1 (Patentin)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/237,624
/ FILING DATE: 10-Sep-2002
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/797,689
/ FILING DATE: 31-JAN-1997
/ APPLICATION NUMBER: US 08/256,927
/ FILING DATE: 28-JUL-1994
/ APPLICATION NUMBER: FR 92/01064
/ FILING DATE: 31-JAN-1992
/ APPLICATION NUMBER: PCT/FR93/00085
/ FILING DATE: 28-JAN-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Smith Ph.D., Julie K.
/ REGISTRATION NUMBER: P-38,619
/ REFERENCE/DOCKET NUMBER: ST92006-US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (610) 454-3839
/ TELEFAX: (610) 454-3808
/ INFORMATION FOR SEQ ID NO: 12:
```

```

;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-624-12

Query Match          100.0%; Score 793; DB 14; Length 138;
Best Local Similarity 100.0%; Pred. No. 3.7e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNEHQVPSNCDLNGGTCVSNKYFSNIHWNCPCPKFGGQHCIEDKSKTCYEGNGHGYRG 60
DB 4 SNEHQVPSNCDLNGGTCVSNKYFSNIHWNCPCPKFGGQHCIEDKSKTCYEGNGHGYRG 63
QY 61 KASTDTMGRPCLPNSATVLAQTYHAHRSALQLGLGKHNYCRNPDNRRPWCYVQVGLK 120
DB 64 KASTDTMGRPCLPNSATVLAQTYHAHRSALQLGLGKHNYCRNPDNRRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
DB 124 PLVQECMVHDCADGK 138

RESULT 8
US-10-702-536-12
; Sequence 12, Application US/10702536
; Publication NO. US20040086976A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guittou, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
;
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
;
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10702,536
; FILING DATE: 07-Nov-2003
; CLASSIFICATION: 435
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3899
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
;
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-702-536-12

Query Match          100.0%; Score 793; DB 16; Length 138;
Best Local Similarity 100.0%; Pred. No. 3.7e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNEHQVPSNCDLNGGTCVSNKYFSNIHWNCPCPKFGGQHCIEDKSKTCYEGNGHGYRG 60
DB 4 SNEHQVPSNCDLNGGTCVSNKYFSNIHWNCPCPKFGGQHCIEDKSKTCYEGNGHGYRG 63
QY 61 KASTDTMGRPCLPNSATVLAQTYHAHRSALQLGLGKHNYCRNPDNRRPWCYVQVGLK 120
DB 64 KASTDTMGRPCLPNSATVLAQTYHAHRSALQLGLGKHNYCRNPDNRRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
DB 124 PLVQECMVHDCADGK 138

RESULT 9
US-10-702-636-12
; Sequence 12, Application US/10702636
; Publication NO. US20040086977A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guittou, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
;
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITIO
; CONTAINING SAID POLYPEPTIDES
;
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10702,636
; FILING DATE: 06-Nov-2003
; CLASSIFICATION: 435
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
;
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-702-536-12
```

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
US-10-702-636-12

Query Match 100.0%; Score 793; DB 16; Length 138;  
Best Local Similarity 100.0%; Pred. No. 3.7e-72;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFGQHCCEIDKSKTCYEGNGHFYRG 60  
DB 4 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFGQHCCEIDKSKTCYEGNGHFYRG 63  
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRPRPCVQVGLK 120  
DB 64 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRPRPCVQVGLK 123  
QY 121 PLVQECMVHDCADGK 135  
DB 124 PLVQECMVHDCADGK 138

## RESULT 10

US-09-880-503-8  
Sequence 8, Application US/09880503  
Patent No. US20020131964A1  
GENERAL INFORMATION:  
APPLICANT: CINES, Douglas B  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND  
FILE REFERENCE: 9596-331  
CURRENT FILING DATE: 2001-06-13  
PRIOR APPLICATION NUMBER: US 60/212,847  
PRIOR FILING DATE: 2000-06-20  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 8  
LENGTH: 143  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-880-503-8

Query Match 100.0%; Score 793; DB 9; Length 143;  
Best Local Similarity 100.0%; Pred. No. 3.9e-72;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFGQHCCEIDKSKTCYEGNGHFYRG 60  
DB 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFGQHCCEIDKSKTCYEGNGHFYRG 60  
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRPRPCVQVGLK 120  
DB 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRPRPCVQVGLK 120  
QY 121 PLVQECMVHDCADGK 135  
DB 121 PLVQECMVHDCADGK 135

## RESULT 11

US-10-106-698-6266  
Sequence 6266, Application US/10106698  
Publication No. US20030109690A1  
GENERAL INFORMATION:  
APPLICANT: Ruben et al.  
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide  
FILE REFERENCE: PA005P1  
CURRENT APPLICATION NUMBER: US/10/106,698  
CURRENT FILING DATE: 2002-03-27  
PRIOR APPLICATION NUMBER: PCT/US00/26524

PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US 60/157,137  
PRIOR FILING DATE: 1999-09-29  
PRIOR APPLICATION NUMBER: US 60/163,280  
PRIOR FILING DATE: 1999-11-03  
NUMBER OF SEQ ID NOS: 8564  
SOFTWARE: Patent In Ver. 3.0  
SEQ ID NO 6266  
LENGTH: 337  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-106-698-6266

Query Match 100.0%; Score 793; DB 14; Length 337;  
Best Local Similarity 100.0%; Pred. No. 9.7e-72;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFGQHCCEIDKSKTCYEGNGHFYRG 60  
DB 27 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFGQHCCEIDKSKTCYEGNGHFYRG 86  
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRPRPCVQVGLK 120  
DB 87 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRPRPCVQVGLK 146  
QY 121 PLVQECMVHDCADGK 135  
DB 147 PLVQECMVHDCADGK 161

## RESULT 12

US-10-264-049-2927  
Sequence 2927, Application US/10264049  
Publication No. US20040005579A1  
GENERAL INFORMATION:  
APPLICANT: Birse et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PA133P1  
CURRENT APPLICATION NUMBER: US/10/264,049  
CURRENT FILING DATE: 2002-10-04  
PRIOR APPLICATION NUMBER: PCT/US01/18569  
PRIOR FILING DATE: 2001-06-07  
PRIOR APPLICATION NUMBER: US 60/209,467  
PRIOR FILING DATE: 2000-06-07  
NUMBER OF SEQ ID NOS: 4360  
SOFTWARE: Patent In Ver. 3.1  
SEQ ID NO 2927  
LENGTH: 337  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-264-049-2927

Query Match 100.0%; Score 793; DB 15; Length 337;  
Best Local Similarity 100.0%; Pred. No. 9.7e-72;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFGQHCCEIDKSKTCYEGNGHFYRG 60  
DB 27 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFGQHCCEIDKSKTCYEGNGHFYRG 86  
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRPRPCVQVGLK 120  
DB 87 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRPRPCVQVGLK 146  
QY 121 PLVQECMVHDCADGK 135  
DB 147 PLVQECMVHDCADGK 161

## RESULT 13

US-09-880-503-6  
Sequence 6, Application US/09880503  
Patent No. US20020131964A1



; PRIOR APPLICATION NUMBER: US 60/328,523  
 ; PRIOR FILING DATE: 2001-10-10  
 ; PRIOR APPLICATION NUMBER: US 60/344,692  
 ; PRIOR FILING DATE: 2001-10-19  
 ; PRIOR APPLICATION NUMBER: US 60/387,292  
 ; PRIOR FILING DATE: 2002-06-07  
 ; PRIOR APPLICATION NUMBER: US 60/391,777  
 ; PRIOR FILING DATE: 2002-06-25  
 ; PRIOR APPLICATION NUMBER: US 60/396,594  
 ; PRIOR FILING DATE: 2002-07-17  
 ; PRIOR APPLICATION NUMBER: US 60/404,249  
 ; PRIOR FILING DATE: 2002-08-16  
 ; PRIOR APPLICATION NUMBER: US 60/407,527  
 ; PRIOR FILING DATE: 2002-08-28  
 ; NUMBER OF SEQ ID NOS: 75  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 34  
 ; LENGTH: 431  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-411-037-34

Query Match 100.0%; Score 793; DB 12; Length 431;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-71;  
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	SNEIHQVPSNCDLGGTGVCSNKYFSNIHWNCNPKKFGQCEIDKSKTCYEGNGHFFYRG	60
Db	21	SNEIHQVPSNCDLGGTGVCSNKYFSNIHWNCNPKKFGQCEIDKSKTCYEGNGHFFYRG	80
QY	61	KASDTDMGRPCLPWNSATVLQCTVHAHRSDALQLGLGKHNYCRNPDNRPPWCYVQVGLK	120
Db	81	KASDTDMGRPCLPWNSATVLQCTVHAHRSDALQLGLGKHNYCRNPDNRPPWCYVQVGLK	140
QY	121	PLVQECMVHDCADGK	135
Db	141	PLVQECMVHDCADGK	155

RESULT 17  
 US-10-411-026-34  
 ; Sequence 34, Application US/10411026  
 ; Publication NO. US20040063911A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Neose Technologies, Inc.  
 ; APPLICANT: DePress, Shawn  
 ; APPLICANT: Zopf, David  
 ; APPLICANT: Bayer, Robert  
 ; APPLICANT: Hakes, David  
 ; APPLICANT: Chen, Xi  
 ; TITLE OF INVENTION: PROTEIN REMODELING METHODS AND PEPTIDES PRODUCED BY THE  
 ; FILE OF INVENTION: METHODS  
 ; FILE REFERENCE: 040853-01-5053  
 ; CURRENT APPLICATION NUMBER: US/10/411,026  
 ; CURRENT FILING DATE: 2003-04-09  
 ; PRIOR APPLICATION NUMBER: US 60/328,523  
 ; PRIOR FILING DATE: 2001-10-10  
 ; PRIOR APPLICATION NUMBER: US 60/344,692  
 ; PRIOR FILING DATE: 2001-10-19  
 ; PRIOR APPLICATION NUMBER: US 60/387,292  
 ; PRIOR FILING DATE: 2002-06-07  
 ; PRIOR APPLICATION NUMBER: US 60/391,777  
 ; PRIOR FILING DATE: 2002-06-25  
 ; PRIOR APPLICATION NUMBER: US 60/396,594  
 ; PRIOR FILING DATE: 2002-07-17  
 ; PRIOR APPLICATION NUMBER: US 60/404,249  
 ; PRIOR FILING DATE: 2002-08-16  
 ; PRIOR APPLICATION NUMBER: US 60/407,527  
 ; PRIOR FILING DATE: 2002-08-28  
 ; NUMBER OF SEQ ID NOS: 75  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 34  
 ; LENGTH: 431

```

; APPLICANT: Hoersch, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 184
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-184

Query Match 100.0%; Score 793; DB 14; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.3e-71;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPVSNCDCCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 21 SNELHQPVSNCDCCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 80
QY 61 KASDTMTGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120
DB 81 KASDTMTGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDRRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK 135
DB 141 PLVQECMVHDCADGK 155

RESULT 20
US-10-193-656-4
; Sequence 4, Application US/10193656
; Publication No. US20030096733A1
; GENERAL INFORMATION:
; APPLICANT: NY, Tor
; APPLICANT: HOLMDAHL, Rikard
; APPLICANT: LI, Jinan
; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
; FILE REFERENCE: 3810/1J577-US3
; CURRENT APPLICATION NUMBER: US/10/193,656
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/304,461
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/304,490
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,182
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank / P00749
; DATABASE ENTRY DATE: 1986-07-21
; RELEVANT RESIDUES: (1)...(431)
US-10-193-656-4

Query Match 100.0%; Score 793; DB 14; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.3e-71;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPVSNCDCCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
US-10-193-656-4
```

```

DB 21 SNELHQPVSNCDCCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 80
QY 61 KASDTMTGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120
DB 81 KASDTMTGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDRRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK 135
DB 141 PLVQECMVHDCADGK 155

RESULT 21
US-10-301-822-161
; Sequence 161, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AN
; FILE REFERENCE: MP01-0292P2RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 161
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-301-822-161

Query Match 100.0%; Score 793; DB 14; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.3e-71;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPVSNCDCCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 21 SNELHQPVSNCDCCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 80
QY 61 KASDTMTGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120
DB 81 KASDTMTGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDRRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK 135
DB 141 PLVQECMVHDCADGK 155

RESULT 22
US-10-131-985-21
; Sequence 21, Application US/10131985
; Publication No. US20030199440A1
; GENERAL INFORMATION:
; APPLICANT: Dack, Kevin N
; APPLICANT: Davies, Michael J
; APPLICANT: Fish, Paul V
; APPLICANT: Huggins, Jonathan P
; APPLICANT: McIntosh, Fraser S
; APPLICANT: Occleston, Nicholas L
US-10-131-985-21
```

```
; TITLE OF INVENTION: Composition
; FILE REFERENCE: PGS 10391A
; CURRENT APPLICATION NUMBER: US/10/131,985
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: US/09/726,295
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: GB 9930768.8
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-985-21

Query Match      100.0%; Score 793; DB 14; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.3e-71;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPVSNCDCGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
   |||||
Db 21 SNELHQPVSNCDCGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 80

QY 61 KASTDTMGRPCPLPWSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120
   |||||
Db 81 KASTDTMGRPCPLPWSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 140

QY 121 PLVQECMVHDCADGK 135
   |||||
Db 141 PLVQECMVHDCADGK 155

RESULT 23
US-10-295-027-414
; Sequence 414, Application US/10295027
; Publication No. US2003023350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1275
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens

; TITLE OF INVENTION: Composition
; FILE REFERENCE: PGS 10391A
; CURRENT APPLICATION NUMBER: US/10/131,985
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: US/09/726,295
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: GB 9930768.8
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-985-21

Query Match      100.0%; Score 793; DB 14; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.3e-71;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPVSNCDCGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
   |||||
Db 21 SNELHQPVSNCDCGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 80

QY 61 KASTDTMGRPCPLPWSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120
   |||||
Db 81 KASTDTMGRPCPLPWSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 140

QY 121 PLVQECMVHDCADGK 135
   |||||
Db 141 PLVQECMVHDCADGK 155

RESULT 23
US-10-295-027-414
; Sequence 414, Application US/10295027
; Publication No. US2003023350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1275
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
```

Query Match	100.0%;	Score 793;	DB 16;	Length 431;
Best Local Similarity	100.0%;	Pred. No. 1.3e-71;		
Matches 135;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
Qy 1	SNELHQVPSNCDCLNGGTCVSNKYFSNTHWNCNPKKFGGQHCIDKSKTCYEGNGHFYRG 60			
Db 21				
Qy 61	KASTDTMGRPCLPWNASATVLQOTYHAHRSDALQLGLGKNYCRPNDRRPPWCYVOVGLK 120			
Db 81	KASTDTMGRPCLPWNASATVLQOTYHAHRSDALQLGLGKNYCRPNDRRPPWCYVOVGLK 140			
Qy 121	PLVQECMVHDCADGK 135			
Db 141	PLVQECMVHDCADGK 155			

```

Query Match      100.0%; Score 793; DB 16; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.3e-71;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY      1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
Db      21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 80

QY      61 KASTDTMGRPCLPWNASATVLOOTYHAHRSDALQIGLGKHNYCRPNDRRSPWCYVOVGLK 120
Db      81 KASTDTMGRPCLPWNASATVLOOTYHAHRSDALQIGLGKHNYCRPNDRRSPWCYVOVGLK 140

QY      121 PLVQECMVHDCADGK 135
Db      141 PLVQECMVHDCADGK 155

RESULT 27
US-10-087-192-594
; Sequence 594, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586

```



```
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 594
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-087-192-594

Query Match      100.0%; Score 793; DB 12; Length 437;
Best Local Similarity 100.0%; Pred. No. 1.3e-71;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHVPNSDCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHGYRG 60
DB 27 SNELHVPNSDCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHGYRG 86
QY 61 KASDTDMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRPRPCVQVGLK 120
DB 87 KASDTDMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRPRPCVQVGLK 146
QY 121 PLVQECMVHDCADGK 135
DB 147 PLVQECMVHDCADGK 161

RESULT 28
US-10-247-671-149
; Sequence 149, Application US/10247671
; Publication No. US20030194721A1
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shifman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247,671
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/323,784
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO 149
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030194721A1 1453334CD1
US-10-247-671-149

Query Match      99.7%; Score 791; DB 14; Length 431;
Best Local Similarity 99.3%; Pred. No. 2e-71;
Matches 134; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHVPNSDCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHGYRG 60
DB 21 SNELHVPNSDCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHGYRG 80
QY 61 KASDTDMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRPRPCVQVGLK 120
DB 81 KASDTDMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRPRPCVQVGLK 140
QY 121 PLVQECMVHDCADGK 135
DB 141 PLVQECMVHDCADGK 155

RESULT 29
US-10-407-821-2
; Sequence 2, Application US/10407821
; Publication No. US20030219386A1
; GENERAL INFORMATION:
; APPLICANT: IDELL, STEVEN
; TITLE OF INVENTION: INTRAPLEURAL SINGLE-CHAIN UROKINASE ALONE OR COMPLEXED
; TITLE OF INVENTION: TO ITS SOLUBLE RECEPTOR PROTECTS AGAINST PLEURAL
; TITLE OF INVENTION: ADHESIONS IN TETRACYCLINE-INDUCED PLEURITIS IN RABBITS
; FILE REFERENCE: USN:022US
; CURRENT APPLICATION NUMBER: US/10/407,821
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 60/414,202
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 60/370,466
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-407-821-2

Query Match      98.7%; Score 783; DB 15; Length 411;
Best Local Similarity 99.3%; Pred. No. 1.2e-70;
Matches 134; Conservative 0; Mismatches 1; Indels 0; Gaps 0

QY 1 SNELHVPNSDCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHGYRG 60
DB 1 SNELHVPNSDCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHGYRG 60
QY 61 KASDTDMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRPRPCVQVGLK 120
DB 61 KASDTDMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRPRPCVQVGLK 120
QY 121 PLVQECMVHDCADGK 135
DB 121 PLVQECMVHDCADGK 135

RESULT 30
US-10-282-174-562
; Sequence 562, Application US/10282174
; Publication No. US20030224380A1
; GENERAL INFORMATION:
; APPLICANT: Becker, Kenneth David
; APPLICANT: Velicelebi, Gonul
; APPLICANT: Eliot, Kathryn J.
; APPLICANT: Wang, Xin
; APPLICANT: Tanzi, Rudolph E.
; APPLICANT: Bertram, Lars
; APPLICANT: Saunders, Aleister J.
; APPLICANT: Mullin, Kristina M.
; APPLICANT: Sampson, Andrew Johnson
; APPLICANT: Blacker, Deborah Lynne
; TITLE OF INVENTION: GENES AND POLYMORPHISMS ON CHROMOSOME 10
; TITLE OF INVENTION: ASSOCIATED WITH ALZHEIMER'S DISEASE AND OTHER
; TITLE OF INVENTION: NEURODEGENERATIVE DISEASES
; FILE REFERENCE: 37481-3308
; CURRENT APPLICATION NUMBER: US/10/282,174
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: US 60/339,525
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 60/338,010
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/336,929
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/338,363
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 60/337,052
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: US 60/368,919
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 564
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 562
```

```
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 15, 58, 141, 214, 231, 274, 366
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-282-174-562

Query Match          98.0%; Score 777; DB 12; Length 431;
Best Local Similarity 98.5%; Pred. No. 5.2e-70;
Matches 133; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SNEHLQVPSNCDLNGGTCVSNKYF--SNHWCNCPKFKGGQHCIDKSKTCYEGNGHGY 60
DB 21 SNEHLQVPSNCDLNGGTCVSNKYF--SNHWCNCPKFKGGQHCIDKSKTCYEGNGHGY 80
QY 61 KASDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRPNDRRPPWCYVQGLK 120
DB 81 KASDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRPNDRRPPWCYVQGLK 140
QY 121 PLVQECMVHDCADGK 135
DB 141 XLVQECMVHDCADGK 155

RESULT 31
US-10-360-101-266
; Sequence 266, Application US/10360101
; Publication No. US20040009550A1
; GENERAL INFORMATION:
; APPLICANT: Moll, Gert N.
; TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way
; FILE REFERENCE: 2183-5673
; CURRENT APPLICATION NUMBER: US/10/360,101
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: EP 02077060.8
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 309
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 266
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence of urokinase
US-10-360-101-266

Query Match          97.2%; Score 771; DB 15; Length 445;
Best Local Similarity 97.1%; Pred. No. 2.2e-69;
Matches 135; Conservative 0; Mismatches 0; Indels 4; Gaps 2;

QY 1 SNEHLQVPSNCDLNGGTCVSNKYF--SNHWCNCPKFKGGQHCIDKSKTCYEGNGHGY 58
DB 21 SNEHLQVPSNCDLNGGTCVSNKYF--SNHWCNCPKFKGGQHCIDKSKTCYEGNGHGY 80
QY 59 RKGASTDTMGRPCLPWNSATVLTQTY--HAHRSDALQLGLGKHNYCRPNDRRPPWCYVQ 116
DB 81 RKGASTDTMGRPCLPWNSATVLTQTY--HAHRSDALQLGLGKHNYCRPNDRRPPWCYVQ 140
QY 117 VGLKPLVQECMVHDCADGK 135
DB 141 VGLKPLVQECMVHDCADGK 159

RESULT 32
US-10-401-077-1
; Sequence 1, Application US/10401077
; Publication No. US20040002137A1
; GENERAL INFORMATION:
; APPLICANT: Hung, Paul Porwen
; APPLICANT: Wu, Bryan T. H.
```

```
; TITLE OF INVENTION: HUMAN TISSUE UROKINASE TYPE PLASMINOGEN
; TITLE OF INVENTION: ACTIVATOR PRODUCTION
; FILE REFERENCE: 12133-006001
; CURRENT APPLICATION NUMBER: US/10/401,077
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/371,013
; PRIOR FILING DATE: 2002-04-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-401-077-1
```

```
Query Match          75.1%; Score 595.5; DB 15; Length 650;
Best Local Similarity 77.9%; Pred. No. 1.6e-51;
Matches 106; Conservative 7; Mismatches 18; Indels 5; Gaps 3

QY 3 ELHQP-SNCD---CLNGGTCVSNKYF--SNHWCNCPKFKGGQHCIDKSKTCYEGNGHGY 58
DB 77 QCHSVFVKSCSEPRFCNGGTCQALYFSDF--VCQCFEGFAGKCCCEIDTRATCYEGNGHGY 135
QY 59 RKGASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRPNDRRPPWCYVQV 118
DB 136 RKGASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKHNYCRPNDRRPPWCYVQV 195
QY 119 LXPLVQECMVHDCADG 134
DB 196 LXPLVQECMVHDCSEG 211
```

```
RESULT 33
US-10-087-192-591
; Sequence 591, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 591
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-087-192-591
```

```
Query Match          69.1%; Score 548; DB 12; Length 433;
Best Local Similarity 71.7%; Pred. No. 6.4e-47;
Matches 91; Conservative 13; Mismatches 23; Indels 0; Gaps 0

QY 9 SNCDLNGGTCVSNKYF--SNHWCNCPKFKGGQHCIDKSKTCYEGNGHGYRKGASTDTMG 68
DB 30 SNGCGQGGVCVSYKYFSRIRRCSCPRFQEGHCEIDASKTCYHONGSDSYRGKANTDTYG 89
QY 69 RFLCPWNSATVLTQTYHAHRSDALQLGLGKHNYCRPNDRRPPWCYVQVGLKPLVQECV 128
DB 90 RFLCPWNSATVLTQTYHAHRSDALQLGLGKHNYCRPNDRRPPWCYVQVGLKPLVQECV 149
QY 129 HDCADGK 135
DB 150 HDCADGK 156
```

RESULT 34  
US-09-880-503-1  
; Sequence 1, Application US/09880503  
; Patent No. US20020131964A1  
; GENERAL INFORMATION:  
; APPLICANT: CINES, Douglas B  
; APPLICANT: HIGAZI, Abd Al-Roof  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND  
; FILE REFERENCE: 9596-331  
; CURRENT APPLICATION NUMBER: US/09/880,503  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/212,847  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 88  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-880-503-1

Query Match 64.3%; Score 510; DB 9; Length 88;  
Best Local Similarity 100.0%; Pred. No. 7.8e-44;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 48 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNCRPN 107  
Db 1 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNCRPN 60  
QY 108 RRRPCYVQVGLKPLVQECVHDCADGK 135  
Db 61 RRRPCYVQVGLKPLVQECVHDCADGK 88

RESULT 35  
US-09-880-503-9  
; Sequence 9, Application US/09880503  
; Patent No. US20020131964A1  
; GENERAL INFORMATION:  
; APPLICANT: CINES, Douglas B  
; APPLICANT: HIGAZI, Abd Al-Roof  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND  
; FILE REFERENCE: 9596-331  
; CURRENT APPLICATION NUMBER: US/09/880,503  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/212,847  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 96  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-880-503-9

Query Match 64.3%; Score 510; DB 9; Length 96;  
Best Local Similarity 100.0%; Pred. No. 8.5e-44;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 48 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNCRPN 107  
Db 1 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNCRPN 60  
QY 108 RRRPCYVQVGLKPLVQECVHDCADGK 135  
Db 61 RRRPCYVQVGLKPLVQECVHDCADGK 88

RESULT 36  
US-10-424-999-37  
; Sequence 37, Application US/10424999

Publication No. US20040052810A1  
; GENERAL INFORMATION:  
; APPLICANT: Nesbit, Mark  
; APPLICANT: Cameron, Beatrice  
; APPLICANT: Blanche, Francis  
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and M  
; FILE REFERENCE: ST01027-A  
; CURRENT APPLICATION NUMBER: US/10/424,999  
; PRIOR FILING DATE: 2003-04-29  
; PRIOR APPLICATION NUMBER: 10/233,675  
; PRIOR FILING DATE: 2002-09-04  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: Patent In version 3.2  
; SEQ ID NO 37  
; LENGTH: 91  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Abrogen D43  
US-10-424-999-37

Query Match 63.2%; Score 501; DB 12; Length 91;  
Best Local Similarity 92.6%; Pred. No. 6.5e-43;  
Matches 88; Conservative 1; Mismatches 2; Indels 4; Gaps 1  
QY 39 GQCEIDKSKTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGK 98  
Db 1 GSH----MAKTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGK 56  
QY 99 HNYCRPNDRRRPWCYVQVGLKPLVQECVHDCAD 133  
Db 57 HNYCRPNDRRRPWCYVQVGLKPLVQECVHDCAD 91

RESULT 37  
US-10-424-999-36  
; Sequence 36, Application US/10424999  
; Publication No. US20040052810A1  
; GENERAL INFORMATION:  
; APPLICANT: Nesbit, Mark  
; APPLICANT: Cameron, Beatrice  
; APPLICANT: Blanche, Francis  
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and M  
; FILE REFERENCE: ST01027-A  
; CURRENT APPLICATION NUMBER: US/10/424,999  
; PRIOR FILING DATE: 2003-04-29  
; PRIOR APPLICATION NUMBER: 10/233,675  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: Patent In version 3.2  
; SEQ ID NO 36  
; LENGTH: 221  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: TrxA-Abrogen D43 fusion protein  
US-10-424-999-36

Query Match 63.2%; Score 501; DB 12; Length 221;  
Best Local Similarity 92.6%; Pred. No. 1.7e-42;  
Matches 88; Conservative 1; Mismatches 2; Indels 4; Gaps 1  
QY 39 GQCEIDKSKTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGK 98  
Db 131 GSH----MAKTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGK 186  
QY 99 HNYCRPNDRRRPWCYVQVGLKPLVQECVHDCAD 133  
Db 187 HNYCRPNDRRRPWCYVQVGLKPLVQECVHDCAD 221

```
RESULT 38
US-10-424-999-17
; Sequence 17, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Using Them to Inhibit Angiogenesis
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 17
; LENGTH: 687
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein human abrogen
US-10-424-999-17
Query Match 63.2%; Score 501; DB 12; Length 687;
Best Local Similarity 97.7%; Pred. No. 5.7e-42;
Matches 86; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 47 SKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKHNCRNPD 106
:|||||
DB 1 AKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKHNCRNPD 60
QY 107 NRRPWCYVQVGLKPLVQECMVHDCADG 134
:|||||
DB 61 NRRPWCYVQVGLKPLVQECMVHDCADG 88
RESULT 39
US-10-425-000-37
; Sequence 37, Application US/10425000
; Publication No. US20040052777A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Kingle Polypeptides and Methods for Using Them to Inhibit
; TITLE OF INVENTION: Angiogenesis
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 37
; LENGTH: 687
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human derived fusion protein
US-10-425-000-37
Query Match 63.2%; Score 501; DB 12; Length 687;
Best Local Similarity 97.7%; Pred. No. 5.7e-42;
Matches 86; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 47 SKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKHNCRNPD 106
:|||||
DB 1 AKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKHNCRNPD 60
QY 107 NRRPWCYVQVGLKPLVQECMVHDCADG 134
:|||||
DB 61 NRRPWCYVQVGLKPLVQECMVHDCADG 88
RESULT 39
US-10-425-000-37
; Sequence 37, Application US/10425000
; Publication No. US20040052777A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Kingle Polypeptides and Methods for Using Them to Inhibit
; TITLE OF INVENTION: Angiogenesis
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 37
; LENGTH: 687
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human derived fusion protein
US-10-425-000-37
Query Match 63.2%; Score 501; DB 12; Length 687;
Best Local Similarity 97.7%; Pred. No. 5.7e-42;
Matches 86; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 47 SKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKHNCRNPD 106
:|||||
DB 1 AKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKHNCRNPD 60
QY 107 NRRPWCYVQVGLKPLVQECMVHDCADG 134
:|||||
DB 61 NRRPWCYVQVGLKPLVQECMVHDCADG 88
```

```
RESULT 38
US-10-424-999-17
; Sequence 17, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
; TITLE OF INVENTION: Using Them to Inhibit Angiogenesis
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; PRIOR FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 17
; LENGTH: 687
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein human abrogen
US-10-424-999-17
Query Match 63.2%; Score 501; DB 12; Length 687;
Best Local Similarity 97.7%; Pred. No. 5.7e-42;
Matches 86; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 47 SKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKHNCRNPD 106
:|||||
DB 1 AKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKHNCRNPD 60
QY 107 NRRPWCYVQVGLKPLVQECMVHDCADG 134
:|||||
DB 61 NRRPWCYVQVGLKPLVQECMVHDCADG 88
RESULT 39
US-10-425-000-37
; Sequence 37, Application US/10425000
; Publication No. US20040052777A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Kingle Polypeptides and Methods for Using Them to Inhibit
; TITLE OF INVENTION: Angiogenesis
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 37
; LENGTH: 687
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human derived fusion protein
US-10-425-000-37
Query Match 63.2%; Score 501; DB 12; Length 687;
Best Local Similarity 97.7%; Pred. No. 5.7e-42;
Matches 86; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 47 SKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKHNCRNPD 106
:|||||
DB 1 AKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKHNCRNPD 60
QY 107 NRRPWCYVQVGLKPLVQECMVHDCADG 134
:|||||
DB 61 NRRPWCYVQVGLKPLVQECMVHDCADG 88
```

```
Db 61 NRRPWCYVQVGLKPLVQECMVHDCADG 88
RESULT 40
US-10-233-675A-17
; Sequence 17, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Fong, Timothy
; APPLICANT: Brockstedt, Dirk
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and M
; TITLE OF INVENTION: Them To Inhibit Angiogenesis
; FILE REFERENCE: ST01027
; CURRENT APPLICATION NUMBER: US/10/233,675A
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/316,300
; PRIOR FILING DATE: 2001-09-04
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 17
; LENGTH: 687
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fusion protein human abrogen
US-10-233-675A-17
Query Match 63.2%; Score 501; DB 15; Length 687;
Best Local Similarity 97.7%; Pred. No. 5.7e-42;
Matches 86; Conservative 2; Mismatches 0; Indels 0; Gaps 0
QY 47 SKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKHNCRNPD 106
:|||||
DB 1 AKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKHNCRNPD 60
QY 107 NRRPWCYVQVGLKPLVQECMVHDCADG 134
:|||||
DB 61 NRRPWCYVQVGLKPLVQECMVHDCADG 88
RESULT 41
US-10-424-999-9
; Sequence 9, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and M
; TITLE OF INVENTION: Using Them to Inhibit Angiogenesis
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424,999
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 9
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human abrogen as secreted from pMB063 (abrogen D43)
US-10-424-999-9
Query Match 63.1%; Score 500; DB 12; Length 87;
Best Local Similarity 98.9%; Pred. No. 7.9e-43;
Matches 86; Conservative 1; Mismatches 0; Indels 0; Gaps 0
QY 47 SKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKHNCRNPD 106
:|||||
DB 1 AKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOQTYHAHRSDALQLGLGKHNCRNPD 60
```

QY 107 NRRPWCYVQVGLKPLVQECMVHDCAD 133  
Db 61 NRRPWCYVQVGLKPLVQECMVHDCAD 87

## RESULT 42

US-10-425-000-29  
; Sequence 29, Application US/10425000  
; Publication No. US20040052777A1  
; GENERAL INFORMATION:  
; APPLICANT: Nesbit, Mark  
; APPLICANT: Cameron, Beatrice  
; APPLICANT: Blanchet, Francis  
; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit  
; FILE REFERENCE: ST01027-B  
; CURRENT APPLICATION NUMBER: US/10/425,000  
; PRIOR FILING DATE: 2003-04-29  
; PRIOR APPLICATION NUMBER: 10/233,675  
; PRIOR FILING DATE: 2002-09-04  
; NUMBER OF SEQ ID NOS: 105  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 29  
; LENGTH: 87  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Human abrogen (D43)  
US-10-425-000-29

Query Match 63.1%; Score 500; DB 12; Length 87;  
Best Local Similarity 98.9%; Pred. No. 7.8e-43;  
Matches 86; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWSATVLOQTYHAHRSALQGLGKHNYCRNPD 106  
Db 1 AKTCYEGNGHFYRGKASTDTMGRPCLPWSATVLOQTYHAHRSALQGLGKHNYCRNPD 60

QY 107 NRRPWCYVQVGLKPLVQECMVHDCAD 133  
Db 61 NRRPWCYVQVGLKPLVQECMVHDCAD 87

## RESULT 43

US-10-233-675A-9  
; Sequence 9, Application US/10233675A  
; Publication No. US20030228298A1  
; GENERAL INFORMATION:  
; APPLICANT: Nesbit, Mark  
; APPLICANT: Fong, Timothy  
; APPLICANT: Brockstedt, Dirk  
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for  
; FILE REFERENCE: ST01027  
; CURRENT APPLICATION NUMBER: US/10/233,675A  
; PRIOR FILING DATE: 2002-09-04  
; PRIOR APPLICATION NUMBER: 60/316,300  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 9  
; LENGTH: 87  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: human derived fusion protein  
US-10-233-675A-9

Query Match 63.1%; Score 500; DB 15; Length 87;  
Best Local Similarity 98.9%; Pred. No. 7.8e-43;  
Matches 86; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWSATVLOQTYHAHRSALQGLGKHNYCRNPD 106  
Db 1 AKTCYEGNGHFYRGKASTDTMGRPCLPWSATVLOQTYHAHRSALQGLGKHNYCRNPD 60  
QY 107 NRRPWCYVQVGLKPLVQECMVHDCAD 133  
Db 61 NRRPWCYVQVGLKPLVQECMVHDCAD 87

## RESULT 44

US-10-424-999-5  
; Sequence 5, Application US/10424999  
; Publication No. US20040052810A1  
; GENERAL INFORMATION:  
; APPLICANT: Nesbit, Mark  
; APPLICANT: Cameron, Beatrice  
; APPLICANT: Blanchet, Francis  
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and M  
; FILE REFERENCE: ST01027-A  
; CURRENT APPLICATION NUMBER: US/10/424,999  
; PRIOR FILING DATE: 2003-04-29  
; PRIOR APPLICATION NUMBER: 10/233,675  
; PRIOR FILING DATE: 2002-09-04  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 5  
; LENGTH: 86  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Human abrogen  
US-10-424-999-5

Query Match 62.9%; Score 499; DB 12; Length 86;  
Best Local Similarity 100.0%; Pred. No. 9.7e-43;  
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 KTCYEGNGHFYRGKASTDTMGRPCLPWSATVLOQTYHAHRSALQGLGKHNYCRNPD 107  
Db 1 KTCYEGNGHFYRGKASTDTMGRPCLPWSATVLOQTYHAHRSALQGLGKHNYCRNPD 60  
QY 108 RRRPWCYVQVGLKPLVQECMVHDCAD 133  
Db 61 RRRPWCYVQVGLKPLVQECMVHDCAD 86

## RESULT 45

US-10-424-999-62  
; Sequence 62, Application US/10424999  
; Publication No. US20040052810A1  
; GENERAL INFORMATION:  
; APPLICANT: Nesbit, Mark  
; APPLICANT: Cameron, Beatrice  
; APPLICANT: Blanchet, Francis  
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and M  
; FILE REFERENCE: ST01027-A  
; CURRENT APPLICATION NUMBER: US/10/424,999  
; PRIOR FILING DATE: 2003-04-29  
; PRIOR APPLICATION NUMBER: 10/233,675  
; PRIOR FILING DATE: 2002-09-04  
; NUMBER OF SEQ ID NOS: 70  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 62  
; LENGTH: 86  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Human kringle domain ATF-Kringle (Abrogen)  
US-10-424-999-62

Query Match 62.9%; Score 499; DB 12; Length 86;

```

Best Local Similarity 100.0%; Pred. No. 9.7e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 KTCYEGNGHFYFGKASTDIMGPCPLPWNSATVLOQTYHAHRS DALQLGLGKHYCRNP DN 107
Db 1 KTCYEGNGHFYFGKASTDIMGPCPLPWNSATVLOQTYHAHRS DALQLGLGKHYCRNP DN 60

QY 108 RRRPWCYVQVGLKPLVQECMWHDCAD 133
Db 61 RRRPWCYVQVGLKPLVQECMWHDCAD 86

```

```

RESULT 46
US-10-425-000-97
; Sequence 97, Application US/10425000
; Publication No. US20040052777A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit
; TITLE OF INVENTION: Angiogenesis
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 97
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human kringle domain ATF-Kringle (Abrogen)
US-10-425-000-97

```

```

Query Match      62.9%; Score 499; DB 12; Length 86;
Best Local Similarity 100.0%; Pred. NO. 9.7e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 48 KTCYEGNGHFYRGKASTDTMGRCPLPNSATVLQQTYHAHRSDALQGLGKINYCRNPDN 107
Db 1 KTCYEGNGHFYRGKASTDTMGRCPLPNSATVLQQTYHAHRSDALQGLGKINYCRNPDN 60

Qy 108 RRRPWCYVQVGLKPLVQECWYHDCAD 133
Db 61 RRRPWCYVQVGLKPLVQECWYHDCAD 86

```

RESULT 47  
 US-10-233-675A-5  
 ; Sequence 5, Application US/10233675A  
 ; Publication No. US20030228298A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Nesbit, Mark  
 ; APPLICANT: Fong, Timothy  
 ; APPLICANT: Brockstedt, Dirk  
 ; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for  
 ; TITLE OF INVENTION: Them to Inhibit Angiogenesis  
 ; FILE REFERENCE: ST01027  
 ; CURRENT APPLICATION NUMBER: US/10/233,675A  
 ; CURRENT FILING DATE: 2002-09-04  
 ; PRIOR APPLICATION NUMBER: 60/316,300  
 ; PRIOR FILING DATE: 2001-09-04  
 ; NUMBER OF SEQ ID NOS: 27  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 5  
 ; LENGTH: 86  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: human derived abrogen

US-10-233-675A-5

Query Match	62.9%	Score 499	DB 15	Length 86
Best Local Similarity	100.0%	Pred. NO. 9.7e-43		
Matches 86	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	48	KTCYEGNGHFYRGKASTDGTMRGCPWPNSATVLTQQTYHAHRSDALQIQLGKKNYCRNP	107	
Db	1	KTCYEGNGHFYRGKASTDGTMRGCPWPNSATVLTQQTYHAHRSDALQIQLGKKNYCRNP	60	
Qy	108	RRRPMCYVQVGLKPLVQECMVHDCAD	133	
Db	61	RRRPMCYVQVGLKPLVQECMVHDCAD	86	

RESULT 48  
US-10-233-675A-22  
; Sequence 22, Application US/10233675A  
; Publication No. US20030228298A1  
; GENERAL INFORMATION:  
; APPLICANT: Nesbit, Mark  
; APPLICANT: Fong, Timothy  
; APPLICANT: Brockstedt, Dirk  
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them  
; TITLE OF INVENTION: Them To Inhibit Angiogenesis  
; FILE REFERENCE: ST01027  
; CURRENT APPLICATION NUMBER: US/10/233,675A  
; CURRENT FILING DATE: 2002-09-04  
; PRIOR APPLICATION NUMBER: 60/316,300  
; PRIOR FILING DATE: 2001-09-04  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 22  
; LENGTH: 86  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: fragment of human urokinase plasminogen activator  
US-10-233-675A-22

```

Query Match      62.9%; Score 499; DB 15; Length 86;
Best Local Similarity 100.0%; Prod.No. 9.7e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 48 KTCYEGNGHFYRGKASTD TMRGCLPWSATVLQQTYYAHRSDALQGLGKHNKYNCRPN 107
Db 1 KTCYEGNGHFYRGKASTD TMRGCLPWSATVLQQTYYAHRSDALQGLGKHNKYNCRPN 60
QY 108 RRRPWCYVQVGLKPLVQECMWHDCA 133
Db 61 RRRPWCYVQVGLKPLVQECMWHDCA 86

```

```

RESULT 49
US-10-424-999-18
; Sequence 18, Application US/10424999
; Publication No. US20040052810A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids
; TITLE OF INVENTION: Using them to Inhibit Angiogenesis
; FILE REFERENCE: ST01027-A
; CURRENT APPLICATION NUMBER: US/10/424, 999
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 688
; TYPE: PRT

```

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein human abrogen
US-10-424-999-18

Query Match          62.8%; Score 498; DB 12; Length 688;
Best Local Similarity 98.9%; Pred. No. 1.2e-41;
Matches 86; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQQTTHAHRSDALQLGLGKKNYCNPD 106
Db 602 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQQTTHAHRSDALQLGLGKKNYCNPD 661

QY 107 NRRRWPCYVQGLKPLVQECWHDCA 133
Db 662 NRRRWPCYVQGLKPLVQECWHDCA 688

RESULT 50
US-10-425-000-38
; Sequence 38, Application US/10425000
; Publication No. US20040052777A1
; GENERAL INFORMATION:
; APPLICANT: Nesbit, Mark
; APPLICANT: Cameron, Beatrice
; APPLICANT: Blanche, Francis
; TITLE OF INVENTION: Kringle Polypeptides and Methods for Using Them to Inhibit
; FILE REFERENCE: ST01027-B
; CURRENT APPLICATION NUMBER: US/10/425,000
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: 10/233,675
; PRIOR FILING DATE: 2002-09-04
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 38
; LENGTH: 688
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human derived fusion protein
US-10-425-000-38

Query Match          62.8%; Score 498; DB 12; Length 688;
Best Local Similarity 98.9%; Pred. No. 1.2e-41;
Matches 86; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQQTTHAHRSDALQLGLGKKNYCNPD 106
Db 602 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQQTTHAHRSDALQLGLGKKNYCNPD 661

QY 107 NRRRWPCYVQGLKPLVQECWHDCA 133
Db 662 NRRRWPCYVQGLKPLVQECWHDCA 688

Search completed: May 25, 2004, 15:03:41
Job time : 25.61 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2004, 14:48:05 ; Search time 9.34706 Seconds  
(without alignments)  
745.636 Million cell updates/sec

Title: US-09-880-503-4  
Perfect score: 793  
Sequence: 1 SNELHQPVSNCCLNGTGV.....QVGLKPLVQECWHDGADGK 135

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 75 summaries

Database : Issued Patents AA.\*

1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	793	100.0	138	2	US-08-797-689-12
2	793	100.0	138	4	US-09-984-186-12
3	793	100.0	200	4	US-09-101-272G-73
4	793	100.0	208	4	US-03-101-272G-98
5	793	100.0	411	1	US-08-087-163-1
6	793	100.0	411	1	US-08-286-748B-18
7	793	100.0	411	1	US-08-153-799-18
8	793	100.0	411	4	US-03-403-736-2
9	793	100.0	430	1	US-07-942-157A-3
10	793	100.0	430	6	5219569-2
11	793	100.0	431	4	US-09-101-272G-1
12	793	100.0	431	6	518829-1
13	788	99.4	194	4	US-09-101-272G-80
14	788	99.4	201	4	US-09-101-272G-96
15	783	98.7	411	3	US-08-181-816-1
16	782	98.6	411	2	US-08-560-098A-48
17	769	97.0	157	3	US-08-142-530B-25
18	765	94.9	432	2	US-08-560-098A-47
19	514	64.8	365	1	US-08-093-741-83
20	514	64.8	365	1	US-08-720-012-83
21	514	64.8	393	2	US-08-560-098A-44
22	514	64.8	393	3	US-08-967-024C-24
23	514	64.8	393	3	US-08-967-024C-25
24	513	64.7	89	4	US-03-101-272G-62
25	335.5	42.3	477	2	US-08-560-098A-51
26	328.5	41.4	527	1	US-07-609-510B-16
27	328.5	41.4	527	2	US-08-811-949-39

28	328.5	41.4	527	5	PCT-US91-01025A-2
29	328.5	41.4	527	6	5185259-8
30	328.5	41.4	527	6	5520913-1
31	328.5	41.4	527	6	5200340-6
32	328.5	41.4	527	2	US-08-811-949-43
33	328.5	41.4	527	2	US-08-560-098A-50
34	328.5	41.4	527	2	US-08-883-795A-38
35	328.5	41.4	527	4	US-09-703-695A-4
36	328.5	41.4	527	6	5185259-3
37	328.5	41.4	527	6	5200340-2
38	328.5	41.4	527	6	5344773-2
39	321.5	40.5	527	6	5244676-5
40	300	37.8	233	3	US-08-438-745-15
41	300	37.8	233	3	US-08-438-745-17
42	300	37.8	233	3	US-09-219-019-15
43	300	37.8	233	3	US-09-219-019-17
44	300	37.8	233	5	PCT-US94-05669A-15
45	300	37.8	233	5	PCT-US94-05669A-17
46	300	37.8	235	3	US-08-438-745-13
47	300	37.8	235	3	US-09-219-019-13
48	300	37.8	235	5	PCT-US94-05669A-13
49	293	36.9	49	2	US-08-747-915-5
50	293	36.9	49	4	US-09-285-783-5
51	288	36.3	48	3	US-09-219-019-22
52	268	33.8	44	2	US-08-747-915-1
53	268	33.8	44	4	US-09-285-783-1
54	260.5	32.8	655	1	US-08-148-918-12
55	260.5	32.8	655	1	US-08-448-937A-12
56	227	28.6	437	2	US-08-811-949-49
57	227	28.6	437	2	US-08-811-949-51
58	227	28.6	437	2	US-08-811-949-55
59	227	28.6	437	2	US-08-811-949-57
60	227	28.6	472	2	US-08-811-949-63
61	221	27.9	356	1	US-08-427-648-8
62	220.5	27.8	326	4	US-09-411-577-3
63	217	27.4	378	4	US-09-553-498-10
64	217	27.4	378	4	US-09-618-869-10
65	215	27.1	347	2	US-08-811-949-1
66	215	27.1	354	2	US-08-811-949-61
67	215	27.1	355	1	US-08-137-116-1
68	215	27.1	355	1	US-08-217-618-1
69	215	27.1	355	1	US-08-427-640-2
70	215	27.1	355	1	US-08-217-617A-1
71	215	27.1	355	1	US-08-217-616-1
72	215	27.1	355	2	US-08-811-949-45
73	215	27.1	355	2	US-08-811-949-47
74	215	27.1	355	2	US-08-811-949-53
75	215	27.1	355	2	US-08-811-949-59

#### ALIGNMENTS

RESULT 1  
US-08-797-689-12  
Sequence 12, Application US/08797689  
Patent No. 5876969  
GENERAL INFORMATION:  
APPLICANT: Fleer, Reinhard  
APPLICANT: Fournier, Alain  
APPLICANT: Guittou, Jean-Dominique  
APPLICANT: Jung, Gerard  
APPLICANT: Yeh, Patrice  
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,  
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION  
CONTAINING SAID POLYPEPTIDES  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rhone-Poulenc Rorer Inc.  
STREET: 500 Arcola Road, 3C43  
CITY: Collegeville  
STATE: PA  
COUNTRY: USA



```

;
;
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; CURRENT APPLICATION DATA:
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; CLASSIFICATION: 435
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-797-689-12

Query Match 100.0%; Score 793; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 4.2e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWCNCPKFGQHCCEIDKSKTCYEGNGHFYRG 60
Db 4 SNELHQPNSCDLNGGTCVSNKYFSNIHWCNCPKFGQHCCEIDKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGPRCLPWSNATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRRPWCYVQVGLK 120
Db 64 KASTDTMGPRCLPWSNATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
Db 124 PLVQECMVHDCADGK 138

RESULT 2
US-09-984-186-12
; Sequence 12, Application US/09984186
; Patent No. 6686179
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guittou, Jean-Dominique
; Jung, Gerard
; Yen, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

;
;
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,186
; FILING DATE: 29-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
; US-09-984-186-12

```

```

Query Match 100.0%; Score 793; DB 4; Length 138;
Best Local Similarity 100.0%; Pred. No. 4.2e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 SNELHQPNSCDLNGGTCVSNKYFSNIHWCNCPKFGQHCCEIDKSKTCYEGNGHFYRG 60
Db 4 SNELHQPNSCDLNGGTCVSNKYFSNIHWCNCPKFGQHCCEIDKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGPRCLPWSNATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRRPWCYVQVGLK 120
Db 64 KASTDTMGPRCLPWSNATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
Db 124 PLVQECMVHDCADGK 138

```

```

RESULT 3
US-09-101-272G-73
; Sequence 73, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: ATF domain of uPA
; US-09-101-272G-73

```

Query Match 100.0%; Score 793; DB 4; Length 200;

Best Local Similarity 100.0%; Pred. No. 6.4e-72;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
Db 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80  
QY 61 KASDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNCRPNDRRPPWCYVQVGLK 120  
Db 81 KASDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNCRPNDRRPPWCYVQVGLK 140  
QY 121 PLVQECMVHDCADGK 135  
Db 141 PLVQECMVHDCADGK 155

RESULT 4  
US-09-101-272G-98  
; Sequence 98, Application US/09101272G  
; Patent No. 6509445  
; GENERAL INFORMATION:  
; APPLICANT: Nissin Food Products Co., Ltd.  
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR  
; FILE REFERENCE: Q50979  
; CURRENT APPLICATION NUMBER: US/09/101.272G  
; PRIOR FILING DATE: 1998-07-08  
; PRIOR APPLICATION NUMBER: JP 1059/1996  
; PRIOR FILING DATE: 1996-01-08  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 98  
; LENGTH: 208  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: ATPH1-ML chimeric protein  
US-09-101-272G-98

Query Match 100.0%; Score 793; DB 4; Length 208;  
Best Local Similarity 100.0%; Pred. No. 6.7e-72;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
Db 2 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 61  
QY 61 KASDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNCRPNDRRPPWCYVQVGLK 120  
Db 62 KASDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNCRPNDRRPPWCYVQVGLK 121  
QY 121 PLVQECMVHDCADGK 135  
Db 122 PLVQECMVHDCADGK 136

RESULT 5  
US-08-087-163-1  
; Sequence 1, Application US/08087163  
; Patent No. 5472692  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jian-Ning  
; APPLICANT: Gurewicz, Victor  
; TITLE OF INVENTION: PRO-UKINASE MUTANTS  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM PS/2 Model 502 or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/087.163  
FILING DATE: 07/02/93  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Fasse, J. Peter  
REGISTRATION NUMBER: 32,983  
REFERENCE/DOCKET NUMBER: 04353/003001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 411  
TYPE: amino acid  
STRANDEDNESS: N/A  
TOPOLOGY: N/A  
US-08-087-163-1

Query Match 100.0%; Score 793; DB 1; Length 411;  
Best Local Similarity 100.0%; Pred. No. 1.4e-71;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0  
QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
Db 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
QY 61 KASDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNCRPNDRRPPWCYVQVGLK 120  
Db 61 KASDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNCRPNDRRPPWCYVQVGLK 120  
QY 121 PLVQECMVHDCADGK 135  
Db 121 PLVQECMVHDCADGK 135

RESULT 6  
US-08-286-748B-18  
; Sequence 18, Application US/08286748B  
; Patent No. 5759542  
; GENERAL INFORMATION:  
; APPLICANT: Victor Gurewicz  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY  
; TITLE OF INVENTION: OF DRUGS BY PLATELETS FOR THE TREATMENT OF  
; TITLE OF INVENTION: CARDIOVASCULAR AND OTHER DISEASES  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 502 or 55SX  
; OPERATING SYSTEM: MS-DOS (Version 5.0)  
; SOFTWARE: WordPerfect (Version 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/286.748B  
; FILING DATE: August 5, 1994  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:

NAME: J. Peter Fasse  
REGISTRATION NUMBER: 32,983  
REFERENCE/DOCKET NUMBER: 04547/013001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 411  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-286-748B-18

Query Match 100.0%; Score 793; DB 1; Length 411;  
Best Local Similarity 100.0%; Pred. No. 1.4e-71;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60  
DB 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60  
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRRRPWCYVQVGLK 120  
DB 61 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRRRPWCYVQVGLK 120  
QY 121 PLVQECMVHDCADGK 135  
DB 121 PLVQECMVHDCADGK 135

RESULT 7  
US-08-153-799-18  
Sequence 18, Application US/08153799  
Patent No. 5766883  
GENERAL INFORMATION:  
APPLICANT: Ballance, David J  
APPLICANT: Goodey, Andrew R  
TITLE OF INVENTION: Polypeptides  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: R Hain Swope, BOC Health Care Inc  
STREET: 100 Mountain Avenue  
CITY: Murray Hill  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07974  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/153,799  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/847975  
FILING DATE: 06-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 8909916.2  
FILING DATE: 29-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB90/00650  
FILING DATE: 26-APR-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/775952  
FILING DATE: 29-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Swope, R Hain  
REGISTRATION NUMBER: 24864  
REFERENCE/DOCKET NUMBER: 92H832

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 665 2400  
TELEFAX: (908) 771 6159  
TELEX: 219484  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 411 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-153-799-18

Query Match 100.0%; Score 793; DB 1; Length 411;  
Best Local Similarity 100.0%; Pred. No. 1.4e-71;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60  
DB 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60  
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRRRPWCYVQVGLK 120  
DB 61 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRRRPWCYVQVGLK 120  
QY 121 PLVQECMVHDCADGK 135  
DB 121 PLVQECMVHDCADGK 135

RESULT 8  
US-09-403-736-2  
Sequence 2, Application US/09403736  
Patent No. 6638502  
GENERAL INFORMATION:  
APPLICANT: Aventis S.A.  
APPLICANT: Li, Hong  
APPLICANT: LU, He  
APPLICANT: GRISCELLI, Frank  
APPLICANT: OPOLO, Paule  
APPLICANT: SORIA, Claudine  
APPLICANT: RAGOT, Thierry  
APPLICANT: LEGRAND, Yves  
APPLICANT: SORIA, Jeanette  
APPLICANT: NABILAT, Christelle  
APPLICANT: PERRICAUD, Michel  
APPLICANT: YEH, Patrice  
TITLE OF INVENTION: Adenovirus-Mediated Intratumoral Delivery Of An Angiogenesis Inhibitor  
FILE REFERENCE: A2778A-US  
CURRENT APPLICATION NUMBER: US/09/403,736  
CURRENT FILING DATE: 1999-10-26  
PRIOR APPLICATION NUMBER: PCT/BP98/02491  
PRIOR FILING DATE: 1998-11-05  
PRIOR APPLICATION NUMBER: 60/044,980  
PRIOR FILING DATE: 1997-04-28  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 2  
LENGTH: 411  
TYPE: PRT  
ORGANISM: humanurokinase  
US-09-403-736-2

Query Match 100.0%; Score 793; DB 4; Length 411;  
Best Local Similarity 100.0%; Pred. No. 1.4e-71;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60  
DB 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60  
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRRRPWCYVQVGLK 120

Db 61 KASTDTMGRPCLPWSNATVLOQTYHAHRSALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120  
QY 121 PLVQECMVHDCADGK 135  
Db 121 PLVQECMVHDCADGK 135

RESULT 9  
US-07-942-157A-3  
; Sequence 3, Application US/07942157A  
; Patent No. 5648253  
; GENERAL INFORMATION:  
; APPLICANT: Wei, Cha-Mer  
; TITLE OF INVENTION: Inhibitor-Resistant Urokinase  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Kilpatrick & Cody  
; STREET: 1100 Peachtree Street Suite 2800  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30309-4530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/942.157A  
; FILING DATE: 19920908  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/631673  
; FILING DATE: 20-DEC-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fabst, Patrea L.  
; REGISTRATION NUMBER: 31,284  
; REFERENCE/DOCKET NUMBER: TS1108Cont.  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (404)815-6508  
; TELEFAX: (404)815-6555  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 430 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEetical: NO  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..19  
; OTHER INFORMATION: /label= peptide  
; OTHER INFORMATION: /note= "WAP signal"  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 198..203  
; OTHER INFORMATION: /label= modified  
; OTHER INFORMATION: /note= "six amino acids deleted in mutant"  
US-07-942-157A-3

Query Match  
Best Local Similarity 100.0%; Score 793; DB 1; Length 430;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
Db 20 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 79  
QY 61 KASTDTMGRPCLPWSNATVLOQTYHAHRSALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120  
Db 80 KASTDTMGRPCLPWSNATVLOQTYHAHRSALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 139  
QY 121 PLVQECMVHDCADGK 135

Db 140 PLVQECMVHDCADGK 154

RESULT 10  
5219569-2  
; Patent No. 5219569  
; APPLICANT: BLABER, MICHAEL; HEYNEKER, HERBERT L.; VEHAR, GORDON A.  
; TITLE OF INVENTION: PROTEASE RESISTANT UROKINASE  
; NUMBER OF SEQUENCES: 6  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/766,858  
; FILING DATE: 16-AUG-1985  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 725,468  
; FILING DATE: 22-APR-1985  
; SEQ ID NO: 2:  
; LENGTH: 430  
5219569-2

Query Match  
Best Local Similarity 100.0%; Score 793; DB 6; Length 430;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80  
QY 61 KASTDTMGRPCLPWSNATVLOQTYHAHRSALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120  
Db 81 KASTDTMGRPCLPWSNATVLOQTYHAHRSALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 140  
QY 121 PLVQECMVHDCADGK 135  
Db 141 PLVQECMVHDCADGK 155

RESULT 11  
US-09-101-272G-1  
; Sequence 1, Application US/09101272G  
; Patent No. 6509445  
; GENERAL INFORMATION:  
; APPLICANT: Nissin Food Products Co., Ltd.  
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR  
; FILE REFERENCE: Q50979  
; CURRENT APPLICATION NUMBER: US/09/101,272G  
; PRIOR FILING DATE: 1998-07-08  
; PRIOR APPLICATION NUMBER: JP 1059/1996  
; PRIOR FILING DATE: 1996-01-08  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 1:  
; LENGTH: 431  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: mat\_peptide  
; LOCATION: (21)..()  
; OTHER INFORMATION:  
; NAME/KEY: misc\_feature  
; LOCATION: (20)..()  
; OTHER INFORMATION: Urokinase-type plasminogen activator (uPA)  
US-09-101-272G-1

Query Match  
Best Local Similarity 100.0%; Score 793; DB 4; Length 431;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80

QY 61 KASTDTMGRPCLPWSATVLOQTYHAHRSALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120  
Db 81 KASTDTMGRPCLPWSATVLOQTYHAHRSALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 140  
QY 121 PLVQECMVHDCADGK 135  
Db 141 PLVQECMVHDCADGK 155

## RESULT 12

5188829-1  
; Patent No. 5188829  
; APPLICANT: KOBAYASHI, YO-ICHI; OMORI, MUNEKI; YAMADA, CHIKAKO  
; TITLE OF INVENTION: RAPIDLY ACTING PROUROKINASE  
; NUMBER OF SEQUENCES: 23  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/340,007  
; FILING DATE: 18-AUG-1998  
; SEQ ID NO: 1:  
; LENGTH: 431  
5188829-1

Query Match 100.0%; Score 793; DB 6; Length 431;  
Best Local Similarity 100.0%; Pred. No. 1.5e-71;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80  
QY 61 KASTDTMGRPCLPWSATVLOQTYHAHRSALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120  
Db 81 KASTDTMGRPCLPWSATVLOQTYHAHRSALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 140  
QY 121 PLVQECMVHDCADGK 135  
Db 141 PLVQECMVHDCADGK 155

## RESULT 13

US-09-101-272G-80  
; Sequence 80, Application US/09101272G  
; Patent No. 6509445  
; GENERAL INFORMATION:  
; APPLICANT: Nissin Food Products Co., Ltd.  
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR  
; FILE REFERENCE: Q50979  
; CURRENT APPLICATION NUMBER: US/09/101,272G  
; PRIOR FILING DATE: 1998-07-08  
; PRIOR APPLICATION NUMBER: JP 1059/1996  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: Patent version 3.1  
; SEQ ID NO 80  
; LENGTH: 194  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: ATFHI chimeric protein  
US-09-101-272G-80

Query Match 99.4%; Score 788; DB 4; Length 194;  
Best Local Similarity 100.0%; Pred. No. 2e-71;  
Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
Db 2 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCIDKSKTCYEGNGHFYRG 61  
QY 61 KASTDTMGRPCLPWSATVLOQTYHAHRSALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120  
Db 62 KASTDTMGRPCLPWSATVLOQTYHAHRSALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 121

QY 121 PLVQECMVHDCADG 134  
Db 122 PLVQECMVHDCADG 135

## RESULT 14

US-09-101-272G-96  
; Sequence 96, Application US/09101272G  
; Patent No. 6509445  
; GENERAL INFORMATION:  
; APPLICANT: Nissin Food Products Co., Ltd.  
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR  
; FILE REFERENCE: Q50979  
; CURRENT APPLICATION NUMBER: US/09/101,272G  
; PRIOR FILING DATE: 1998-07-08  
; PRIOR APPLICATION NUMBER: JP 1059/1996  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: Patent version 3.1  
; SEQ ID NO 96  
; LENGTH: 201  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: ATFHI-CL chimeric protein  
US-09-101-272G-96

Query Match 99.4%; Score 788; DB 4; Length 201;  
Best Local Similarity 100.0%; Pred. No. 2e-71;  
Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
Db 2 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCIDKSKTCYEGNGHFYRG 61  
QY 61 KASTDTMGRPCLPWSATVLOQTYHAHRSALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120  
Db 62 KASTDTMGRPCLPWSATVLOQTYHAHRSALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 121  
QY 121 PLVQECMVHDCADG 134  
Db 122 PLVQECMVHDCADG 135

## RESULT 15

US-09-181-816-1  
; Sequence 1, Application US/09181816  
; Patent No. 627818  
; GENERAL INFORMATION:  
; APPLICANT: MAZAR, Andrew P.  
; APPLICANT: JONES, Terence R.  
; TITLE OF INVENTION: CYCLIC PEPTIDE LIGANDS THAT TARGET UROKINASE  
; TITLE OF INVENTION: PLASMINOGEN ACTIVATOR RECEPTOR  
; FILE REFERENCE: 32904200300 SIDN 1-7  
; CURRENT APPLICATION NUMBER: US/09/181,816  
; CURRENT FILING DATE: 1998-10-29  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 411  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-181-816-1

Query Match 98.7%; Score 783; DB 3; Length 411;  
Best Local Similarity 99.3%; Pred. No. 1.4e-70;  
Matches 134; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
QY 61 KASTDTMGRPCLPWSATVLOQTYHAHRSALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120

Db 61 KASDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120  
QY 121 PLVQECMVHDCADGK 135  
Db 121 LLVQECMVHDCADGK 135

## RESULT 16

US-08-560-098A-48  
; Sequence 48, Application US/08560098A  
; Patent No. 5976841  
; GENERAL INFORMATION:  
; APPLICANT: WENDEL, Stephan  
; APPLICANT: HEINZEL-WIELAND, Regina  
; APPLICANT: STEFFENS, Gerd Josef  
; TITLE OF INVENTION: Proteins having Fibrinolytic and  
; TITLE OF INVENTION: Coagulation-inhibiting Properties  
; NUMBER OF SEQUENCES: 60  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan  
; STREET: 1200 G Street, N.W., Suite 700  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30 (BPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/560,098A  
; FILING DATE: 17-NOV-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: P 44 40 892.7  
; FILING DATE: 17-NOV-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: EVANS, Joseph D.  
; REGISTRATION NUMBER: 26,269  
; REFERENCE/DOCKET NUMBER: 148/42448  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 628-8800  
; TELEFAX: (202) 628-8844  
; INFORMATION FOR SEQ ID NO: 48:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 411 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-560-098A-48

Query Match 98.6%; Score 782; DB 2; Length 411;  
Best Local Similarity 99.3%; Pred. No. 1.8e-70;  
Matches 134; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60  
Db 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60  
QY 61 KASDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120  
Db 61 KASDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120  
QY 121 PLVQECMVHDCADGK 135  
Db 121 PLVQECMVHDCADGK 135

## RESULT 17

US-08-142-590B-25  
; Sequence 25, Application US/08142590B

; Patent No. 6120765  
; GENERAL INFORMATION:  
; APPLICANT: HIBINO, Tashihiko; TAKAHASHI, Tadahito; HORII, Izumi; and C FINCK,  
; TITLE OF INVENTION: UROKINASE PLASMINOGEN ACTIVATOR FRAGMENTS  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD, LLP  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/142,590B  
; FILING DATE: 25-OCT-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/042,318  
; FILING DATE: 02-APR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Myers, Paul L.  
; REGISTRATION NUMBER: 35,965  
; REFERENCE/DOCKET NUMBER: MGP-009CP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 157 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
US-08-142-590B-25

Query Match 97.0%; Score 769; DB 3; Length 157;  
Best Local Similarity 97.0%; Pred. No. 1.2e-69;  
Matches 131; Conservative 0; Mismatches 4; Indels 0; Gaps 0  
QY 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60  
Db 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60  
QY 61 KASDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120  
Db 61 KASDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120  
QY 121 PLVQECMVHDCADGK 135  
Db 121 PLVQECMVHDCADGK 135

## RESULT 18

US-08-560-098A-47  
; Sequence 47, Application US/08560098A  
; Patent No. 5976841  
; GENERAL INFORMATION:  
; APPLICANT: WENDEL, Stephan  
; APPLICANT: HEINZEL-WIELAND, Regina  
; APPLICANT: STEFFENS, Gerd Josef  
; TITLE OF INVENTION: Proteins having Fibrinolytic and  
; TITLE OF INVENTION: Coagulation-inhibiting Properties  
; NUMBER OF SEQUENCES: 60  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan  
; STREET: 1200 G Street, N.W., Suite 700  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA

```

; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8944
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-560-098A-47

Query Match 64.9%; Score 515; DB 2; Length 432;
Best Local Similarity 75.6%; Pred. No. 1e-43;
Matches 93; Conservative 8; Mismatches 16; Indels 6; Gaps 1;

QY 13 CLNGGTCVSNKYFSNIHNCNCPKFGGCHCEIDKSKTCYEGNGHFGYRGKASTDTWGRPCL 72
Db 40 CVTGGTYPKPSHHNGDPEETPEY-----LQISKTCYEGNGHFGYRGKASTDTWGRPCL 93

QY 73 PWSNATVLQOQTYVHAHRSDALQLGLGKNYCRNPNRRRRPWCYVQVGLKELVQECMVHPCA 132
Db 94 PWSNATVLQOQTYVHAHRSDALQLGLGKNYCRNPNRRRRPWCYVQVGLKELVQECMVHPCA 153

QY 133 DGK 135
Db 154 DGK 156

RESULT 19
US-08-093-741-83
; Sequence 83, Application US/08093741
; Patent No: 5691721
; GENERAL INFORMATION:
; APPLICANT: STEFFENS, GERD J.
; APPLICANT: WRENDT, STEPHAN
; APPLICANT: SCHNEIDER, JOHANNES
; APPLICANT: HEINZEL-WIELAND, REGINA
; APPLICANT: SAUNDERS, DEREK J.
; TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
; TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
; TITLE OF INVENTION: INHIBITING EFFECT
; NUMBER OF SEQUENCES: 83
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N. W. Suite 700
; CITY: Washington, D.C.
; COUNTRY: U.S.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/093,741
; FILING DATE: 20-JUL-1993

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-720-012-83

Query Match 64.8%; Score 514; DB 1; Length 365;
Best Local Similarity 100.0%; Pred. No. 1.1e-43;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNCRNPD 106
Db 1 SKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNCRNPD 60

QY 107 NRRPWCYVQVGLKPLVQECMVHDCADGK 135
Db 61 NRRPWCYVQVGLKPLVQECMVHDCADGK 89

RESULT 21
US-08-560-098A-44
; Sequence 44, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNENDT, Stephan
; APPLICANT: STEFFENS, Gerd Josef
; APPLICANT: HEINZEL-WIELAND, Regina
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-44

Query Match 64.8%; Score 514; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 1.2e-43;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNCRNPD 106
Db 2 SKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNCRNPD 61

QY 107 NRRPWCYVQVGLKPLVQECMVHDCADGK 135
```

```

Db 62 NRRPWCYVQVGLKPLVQECMVHDCADGK 90

RESULT 22
US-08-967-024C-24
; Sequence 24, Application US/08967024C
; Patent No. 6133011
; GENERAL INFORMATION:
; APPLICANT: WNENDT, Stephan
; APPLICANT: STEFFENS, Gerd Josef
; APPLICANT: JANOSHA, Elke
; APPLICANT: HEINZEL-WIELAND, Regina
; TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967,024C
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 42 665.8
; FILING DATE: 30-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42444
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-967-024C-24

Query Match 64.8%; Score 514; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 1.2e-43;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNCRNPD 106
Db 2 SKTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNCRNPD 61

QY 107 NRRPWCYVQVGLKPLVQECMVHDCADGK 135
Db 62 NRRPWCYVQVGLKPLVQECMVHDCADGK 90

RESULT 23
US-08-967-024C-25
; Sequence 25, Application US/08967024C
; Patent No. 6133011
; GENERAL INFORMATION:
; APPLICANT: WNENDT, Stephan
; APPLICANT: STEFFENS, Gerd Josef
; APPLICANT: JANOSHA, Elke
; APPLICANT: HEINZEL-WIELAND, Regina
; TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
; NUMBER OF SEQUENCES: 25
```



```

CORRESPONDENCE ADDRESS:
ADDRESSES: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967.024C
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 42 665.8
FILING DATE: 30-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42444
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-967-024C-25

Query Match 64.8%; Score 514; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 1.2e-43;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGRCLPWNSTVLTQOYTHAHRSDALQGLGKHNYPD 106
DB 2 SKTCYEGNGHFYRGKASTDTMGRCLPWNSTVLTQOYTHAHRSDALQGLGKHNYPD 61

QY 107 NRRPWCYVQVGLKPLVQECMVHDCADGK 135
DB 62 NRRPWCYVQVGLKPLVQECMVHDCADGK 90

RESULT 24
US-09-101-272G-62
Sequence 62, Application US/09101272G
Patent No. 6509445
GENERAL INFORMATION:
APPLICANT: Nissin Food Products Co., Ltd.
TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
FILE REFERENCE: Q50979
CURRENT APPLICATION NUMBER: US/09/101.272G
CURRENT FILING DATE: 1998-07-08
PRIOR APPLICATION NUMBER: JP 1059/1996
PRIOR FILING DATE: 1996-01-08
NUMBER OF SEQ ID NOS: 107
SOFTWARE: Patent in version 3.1
SEQ ID NO 62
LENGTH: 89
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: residues 43-131 of the ATF domain of uPA
US-09-101-272G-62

Query Match 64.7%; Score 513; DB 4; Length 89;
Best Local Similarity 100.0%; Pred. No. 2.7e-44;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CORRESPONDENCE ADDRESS:
ADDRESSES: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967.024C
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 42 665.8
FILING DATE: 30-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42444
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-967-024C-25

Query Match 42.3%; Score 335.5; DB 2; Length 477;
Best Local Similarity 46.3%; Pred. No. 1.1e-25;
Matches 62; Conservative 17; Mismatches 50; Indels 5; Gaps 2

QY 3 ELHQVP----SNCDCLNGGTCVSNKYPFNIHWCNPKKFGQGHCEIDKSKTCYEGNGHFY 58
DB 78 OCHTVPVKSCSELRCFNGTCWQAASDF--VCQCPKGYTGKQCEVDTHATCYKQDQVTV 136

QY 59 RKGASTDTMGRCLPWNSTVLTQOYTHAHRSDALQGLGKHNYPDNRPRPWCYVQVG 116
DB 137 RGTWSTSSGACQCNWNSNLLTRTYNGRRSDAITLGLGNHNYCRNPDDNSKPKWCYVKA 196

QY 119 LKPLVQECMVHDC 132
DB 197 SKFILEFCVPCVS 210

RESULT 26
```

```
US-07-609-510B-16
; Sequence 16, Application US/07609510B
; Patent No. 5326700
; GENERAL INFORMATION:
; APPLICANT: Berg et al.
; TITLE OF INVENTION: Method for Altering Post-Translational Processing of Tissue Pl
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN.
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07609,510B
; FILING DATE: 19901106
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-609-510B-16

Query Match 41.4%; Score 328.5; DB 1; Length 527;
Best Local Similarity 46.3%; Pred. No. 6.1e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHVP-SNCD--CLNGGTCVSNKYFNHWCNCPKFGGCHCEIDKSKTCYEGNGHFY 58
DB 42 QCHSVPVKSCSEPRCFNGTCQOALYFSDF-VCCPEGFAGKCCIDTRATCYEDQGISY 100
QY 59 RGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVG 118
DB 101 RGTWSTAESGAECTNWNSSALAQKPYSGRRPDAIRLGLGNHNYCRNPNDRSKPWCYVFKA 160
QY 119 LKPLVQECMVHDCADG 134
DB 161 GKSSEFCSTPACSEG 176

RESULT 27
US-08-811-949-39
; Sequence 39, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: KAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

US-07-609-510B-16
; Sequence 16, Application US/07609510B
; Patent No. 5326700
; GENERAL INFORMATION:
; APPLICANT: Berg et al.
; TITLE OF INVENTION: Method for Altering Post-Translational Processing of Tissue Pl
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN.
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07609,510B
; FILING DATE: 19901106
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-609-510B-16

Query Match 41.4%; Score 328.5; DB 1; Length 527;
Best Local Similarity 46.3%; Pred. No. 6.1e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHVP-SNCD--CLNGGTCVSNKYFNHWCNCPKFGGCHCEIDKSKTCYEGNGHFY 58
DB 42 QCHSVPVKSCSEPRCFNGTCQOALYFSDF-VCCPEGFAGKCCIDTRATCYEDQGISY 100
QY 59 RGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVG 118
DB 101 RGTWSTAESGAECTNWNSSALAQKPYSGRRPDAIRLGLGNHNYCRNPNDRSKPWCYVFKA 160
QY 119 LKPLVQECMVHDCADG 134
DB 161 GKSSEFCSTPACSEG 176

RESULT 27
US-08-811-949-39
; Sequence 39, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: KAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-39

Query Match 41.4%; Score 328.5; DB 2; Length 527;
Best Local Similarity 46.3%; Pred. No. 6.1e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHVP-SNCD--CLNGGTCVSNKYFNHWCNCPKFGGCHCEIDKSKTCYEGNGHFY 58
DB 42 QCHSVPVKSCSEPRCFNGTCQOALYFSDF-VCCPEGFAGKCCIDTRATCYEDQGISY 100
QY 59 RGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVG 118
DB 101 RGTWSTAESGAECTNWNSSALAQKPYSGRRPDAIRLGLGNHNYCRNPNDRSKPWCYVFKA 160
QY 119 LKPLVQECMVHDCADG 134
DB 161 GKSSEFCSTPACSEG 176

RESULT 28
PCT-US91-01025A-2
; Sequence 2, Application PC/TUS9101025A
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: Tissue Plasminogen Activator Having Fibrin
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/01025A
; FILING DATE: 19910214
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/486,657
; FILING DATE: 1 March 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 454P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/266-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
```

```
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
PCT-US91-01025A-2

Query Match          41.4%; Score 328.5; DB 5; Length 527;
Best Local Similarity 46.3%; Pred. No. 6.1e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDSKTKCYEGNGHFY 58
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 59 RKGASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKKNYCRNPNRRPWCYVQVG 118
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 101 RGTWSTAESGAECTWNSSALAQKPYSGRRPDALRLGLGNHNYCRNPNRRPWCYVQVG 160
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 119 LKPLVQECMVHDCADG 134
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
161 GYSSFCSTPACSEG 176

RESULT 29
5185259-8
; Patent No. 5185259
; APPLICANT: GORDELL, DAVID V.; KOHR, WILLIAM J.; PENNICA, DIANE;
; VEHAR, GORDON A.
; TITLE OF INVENTION: TRUNCATED HUMAN TISSUE PLASMINOGEN
; ACTIVATOR
; NUMBER OF SEQUENCES: 15
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/489,855
; FILING DATE: 02-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 12,694
; FILING DATE: 09-FEB-1987
; APPLICATION NUMBER: 483,052
; FILING DATE: 07-APR-1983
; APPLICATION NUMBER: 398,003
; FILING DATE: 14-JUL-1982
; APPLICATION NUMBER: 374,860
; FILING DATE: 05-MAY-1982
; SEQ ID NO: 8.
; LENGTH: 527
5185259-8

Query Match          41.4%; Score 328.5; DB 6; Length 527;
Best Local Similarity 46.3%; Pred. No. 6.1e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDSKTKCYEGNGHFY 58
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 59 RKGASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKKNYCRNPNRRPWCYVQVG 118
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 101 RGTWSTAESGAECTWNSSALAQKPYSGRRPDALRLGLGNHNYCRNPNRRPWCYVQVG 160
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 119 LKPLVQECMVHDCADG 134
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
161 GYSSFCSTPACSEG 176

RESULT 30
5520913-1
; Patent No. 5520913
; APPLICANT: ANDERSON, STEPHEN; BENNETT, WILLIAM F.; BOTSTEIN,
; DAVID; HIGGINS, DEBORAH L.; PAONI, NICHOLAS F.; ZOLLER, MARK J.
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR HAVING
; ZYMOGENIC PROPERTIES
; NUMBER OF SEQUENCES: 35

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
PCT-US91-01025A-2

Query Match          41.4%; Score 328.5; DB 5; Length 527;
Best Local Similarity 46.3%; Pred. No. 6.1e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDSKTKCYEGNGHFY 58
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 59 RKGASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKKNYCRNPNRRPWCYVQVG 118
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 101 RGTWSTAESGAECTWNSSALAQKPYSGRRPDALRLGLGNHNYCRNPNRRPWCYVQVG 160
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 119 LKPLVQECMVHDCADG 134
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
161 GYSSFCSTPACSEG 176

RESULT 31
5200340-6
; Patent No. 5200340
; APPLICANT: POSTER, DONALD C.; MULVIHILL, EILEEN R.; O'HARA,
; PATRICK J.; PINGEL, KURT; YOSHITAKE, SHINJI
; TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN
; ACTIVATORS
; NUMBER OF SEQUENCES: 34
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/53,412
; FILING DATE: 22-MAY-1987
; SEQ ID NO: 6.
; LENGTH: 546
5200340-6

Query Match          41.4%; Score 328.5; DB 6; Length 546;
Best Local Similarity 46.3%; Pred. No. 6.3e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDSKTKCYEGNGHFY 58
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 59 RKGASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKKNYCRNPNRRPWCYVQVG 118
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 136 RGTWSTAESGAECTWNSSALAQKPYSGRRPDALRLGLGNHNYCRNPNRRPWCYVQVG 195
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 119 LKPLVQECMVHDCADG 134
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
196 GYSSFCSTPACSEG 211

RESULT 32
US-08-811-949-43
; Sequence 43, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
```

```

; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 562 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-811-949-43

Query Match 41.4%; Score 328.5; DB 2; Length 562;
Best Local Similarity 46.3%; Pred. No. 6.5e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHGY 58
Db 77 QCHSVFVKSCBPRCFNGTCQQALYFSDF-VCQCPEGFAGKCEIDTRATCYEDQGISY 135

QY 59 RGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRNPNRRRRCPCYQVG 118
Db 136 RGTWSTAESGAECTNWNSSALAKPYSGRRPDATRLGLGNHNYCRNPDSDSKPCYVFK 195

QY 119 LKPLVQECMVHDCADG 134
Db 196 GKYSEFCSTPACSEG 211

RESULT 33
US-08-560-098A-50
; Sequence 50, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WENNDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 562 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-811-949-43

Query Match 41.4%; Score 328.5; DB 2; Length 562;
Best Local Similarity 46.3%; Pred. No. 6.5e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHGY 58
Db 77 QCHSVFVKSCBPRCFNGTCQQALYFSDF-VCQCPEGFAGKCEIDTRATCYEDQGISY 135

QY 59 RGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRNPNRRRRCPCYQVG 118
Db 136 RGTWSTAESGAECTNWNSSALAKPYSGRRPDATRLGLGNHNYCRNPDSDSKPCYVFK 195

QY 119 LKPLVQECMVHDCADG 134
Db 196 GKYSEFCSTPACSEG 211

RESULT 34
US-08-883-795A-38
; Sequence 38, Application US/08883795A
; Patent No. 5985607
; GENERAL INFORMATION:
; APPLICANT: Delcuve, Genevieve
; APPLICANT: Awang, Gregor
; TITLE OF INVENTION: Recombinant DNA Molecules and Expression
; TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERESKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/883,795A
; FILING DATE: 27-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gravelle, Micheline
; REGISTRATION NUMBER: 40,261
; REFERENCE/DOCKET NUMBER: 7841-062
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311

```

```

; FILING DATE: 02-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 12,694
; FILING DATE: 09-FEB-1987
; APPLICATION NUMBER: 483,052
; FILING DATE: 07-APR-1983
; APPLICATION NUMBER: 398,003
; FILING DATE: 14-JUL-1982
; APPLICATION NUMBER: 374,860
; FILING DATE: 05-MAY-1982
; SEQ ID NO:3
; LENGTH: 562
5185259-3

Query Match 41.4%; Score 328.5; DB 6; Length 562;
Best Local Similarity 46.3%; Pred. No. 6.5e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3

QY 3 ELHQP-SNCD---CLNGGTCVKNKYPFSNIHWCNCPKKGQGCETDKSKTCYEGNGHFY 58
Db 77 QCHSPVPKSCSEPCFCNGGTCQQALYPSDF-VQCPEGFAGKCEIDTRATCYEDOGISY 135
QY 59 RGKASTDTMGRCPLPWNASATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRRRPWCYVQVG 118
Db 136 RGTWSTAESGAECTNWNSSALAQKPYSGRRPDAIRLGLGNHNYCRNPDNRDRSKPCYVFK 195
QY 119 LKPLVQECMVHDCADG 134
Db 196 GKYSSEFCSTPACSEG 211

RESULT 37
5200340-2
; Patent No. 5200340
; APPLICANT: FOSTER, DONALD C.; MULVIHILL, EILEEN R.; O'HARA,
; PATRICK J.; PINGEL, KURT; YOSHITAKE, SHINJI
; TITLE OF INVENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN
; ACTIVATORS
; NUMBER OF SEQUENCES: 34
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/53,412
; FILING DATE: 22-MAY-1987
; SEQ ID NO:2
; LENGTH: 562
5200340-2

Query Match 41.4%; Score 328.5; DB 6; Length 562;
Best Local Similarity 46.3%; Pred. No. 6.5e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3

QY 3 ELHQP-SNCD---CLNGGTCVKNKYESNIHWCNCPKKGQGCETDKSKTCYEGNGHFY 58
Db 77 QCHSPVPKSCSEPCFCNGGTCQQALYPSDF-VQCPEGFAGKCEIDTRATCYEDOGISY 135
QY 59 RGKASTDTMGRCPLPWNASATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRRRPWCYVQVG 118
Db 136 RGTWSTAESGAECTNWNSSALAQKPYSGRRPDAIRLGLGNHNYCRNPDNRDRSKPCYVFK 195
QY 119 LKPLVQECMVHDCADG 134
Db 196 GKYSSEFCSTPACSEG 211

RESULT 38
5344773-2
; Patent No. 5344773
; APPLICANT: WEI, CHA-WER, HSUNG, NANCY; REDDY, VERMURI B.;
; LEMONITT, JEFFREY F.; DACKOWSKI, WILLIAM; DOUGLAS, RICHARD;
; COLE, EDWARD S.; PURCELL JR., RICHARD D.; LAU, DAVID TAI-YUI
; TITLE OF INVENTION: HUMAN UTERINE TISSUE PLASMINOGEN
; ACTIVATOR PRODUCED BY RECOMBIANT DNA
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/07/782,686  
FILING DATE: 01-OCT-1985  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 656,770  
FILING DATE: 01-OCT-1984  
SEQ ID NO: 2  
LENGTH: 562  
5344773-2

Query Match  
Best Local Similarity 41.4%; Score 328.5; DB 6; Length 562;  
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHQVP-SNCD---CLNGTCTVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFF 58  
DB 77 QCHSVPVKCSBPCFNGTCCQALYFSDP-VQCPGEPGAGKCCIDTRACYEDQGGLSY 135

QY 59 RGAASDTMGRPCLPNNSATVLTQTYHAHRSDALQGLGKHNYCRNPDRRPPWCYQVQG 118  
DB 136 RGTWSTAESGAECTNWNSSALAKQPYSGRRPDALRLGLNHNYCRNPDRDSKPCWYVFK 195

QY 119 LKPLVQECMVHDCADG 134  
DB 196 GRYSSEFCSTPACSEG 211

RESULT 39  
5244676-5  
Patent No. 5244676  
APPLICANT: BELL, LESLIE D.; WAYER, ERNEST J.; PALMIER, MARK O.  
TOLUNAY, H.ESER; WARREN, THOMAS G.; WUN, TZE-CHEIN  
TITLE OF INVENTION: MODIFIED TISSUE PLASMINOGEN ACTIVATOR  
WITH MODIFIED GLYCOSYLATION SITE  
NUMBER OF SEQUENCES: 9  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/203,047  
FILING DATE: 06-JUN-1988  
SEQ ID NO: 5  
LENGTH: 562  
5244676-5

Query Match  
Best Local Similarity 40.5%; Score 321.5; DB 6; Length 562;  
Matches 61; Conservative 16; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHQVP-SNCD---CLNGTCTVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFF 58  
DB 77 QCHSVPVKCSBPCFNGTCCQALYFSDP-VQCPGEPGAGKCCIDGNSDCYFGSGSAY 135

QY 59 RGAASDTMGRPCLPNNSATVLTQTYHAHRSDALQGLGKHNYCRNPDRRPPWCYQVQG 118  
DB 136 RGTWSTAESGAECTNWNSSALAKQPYSGRRPDALRLGLNHNYCRNPDRDSKPCWYVFK 195

QY 119 LKPLVQECMVHDCADG 134  
DB 196 RLWTEYCDVPSCEG 211

RESULT 40  
US-08-438-745-15  
Sequence 15, Application US/08438745  
Patent No. 6248715  
GENERAL INFORMATION:  
APPLICANT: Rosenberg, Steven  
TITLE OF INVENTION: Expression of Urokinase Plasminogen  
Activator Inhibitors  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Chiron Corporation  
STREET: 4560 Horton Street  
CITY: Emeryville  
STATE: CA

COUNTRY: USA  
ZIP: 94608  
COMPUTER READABLE FORM: disk  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/438,745  
FILING DATE: 10-MAY-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/070,153  
FILING DATE: 01-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Green, Grant D.  
REGISTRATION NUMBER: 31,259  
REFERENCE/DOCKET NUMBER: 0939.001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 510-601-2706  
TELEFAX: 510-655-3542  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 233 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-438-745-15

Query Match  
Best Local Similarity 37.8%; Score 300; DB 3; Length 233;  
Matches 53; Conservative 2; Mismatches 13; Indels 0; Gaps 0

QY 1 SNELHQVPSNCDCLNGTCTVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFF 60  
DB 29 SNELHQVPSNCDCLNGTCTVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFF 88

QY 61 KASTDTMG 68  
DB 89 KXANANKG 96

RESULT 41  
US-08-438-745-17  
Sequence 17, Application US/08438745  
Patent No. 6248715  
GENERAL INFORMATION:  
APPLICANT: Rosenberg, Steven  
STRATTON-THOMAS, JENNIFER  
TITLE OF INVENTION: Expression of Urokinase Plasminogen  
Activator Inhibitors  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Chiron Corporation  
STREET: 4560 Horton Street  
CITY: Emeryville  
STATE: CA  
COUNTRY: USA  
ZIP: 94608  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/438,745  
FILING DATE: 10-MAY-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/070,153  
FILING DATE: 01-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Green, Grant D.

```
;
;
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 0939,001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2706
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 233 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-438-745-17

Query Match          37.8%; Score 300; DB 3; Length 233;
Best Local Similarity 77.9%; Pred. No. 1.7e-22;
Matches 53; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIDKSKTCYEGNGHFYRG 60
DB 29 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIDKSKQGGSGGDFDYE 88

QY 61 KASTDTMG 68
DB 89 KMANANKG 96

RESULT 42
US-09-219-019-15
; Sequence 15, Application US/09219019
; Patent No. 6268341
; GENERAL INFORMATION:
; APPLICANT: ROSENBERG, STEVEN
; TITLE OF INVENTION: EXPRESSION OF UROKINASE PLASMINOGEN ACTIVATOR
; TITLE OF INVENTION: INHIBITORS
; FILE REFERENCE: 23533-0005
; CURRENT APPLICATION NUMBER: US/09/219,019
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 08/438,263
; PRIOR FILING DATE: 1995-05-10
; PRIOR APPLICATION NUMBER: 08/280,288
; PRIOR FILING DATE: 1994-07-26
; PRIOR APPLICATION NUMBER: 08/070,153
; PRIOR FILING DATE: 1993-06-01
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 15
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-219-019-15

Query Match          37.8%; Score 300; DB 3; Length 233;
Best Local Similarity 77.9%; Pred. No. 1.7e-22;
Matches 53; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIDKSKTCYEGNGHFYRG 60
DB 29 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIDKSKQGGSGGDFDYE 88

QY 61 KASTDTMG 68
DB 89 KMANANKG 96

RESULT 44
PCT-US94-05669A-15
; Sequence 15, Application PC/TUS9405669A
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; TITLE OF INVENTION: Expression of Urokinase Plasminogen
; TITLE OF INVENTION: Activator Inhibitors
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/05669A
; FILING DATE: 19-MAY-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D.
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 0939,100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2706
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 233 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US94-05669A-15

Query Match          37.8%; Score 300; DB 5; Length 233;
Best Local Similarity 77.9%; Pred. No. 1.7e-22;
Matches 53; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

;
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 0939,001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2706
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 233 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-438-745-17

Query Match          37.8%; Score 300; DB 3; Length 233;
Best Local Similarity 77.9%; Pred. No. 1.7e-22;
Matches 53; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIDKSKTCYEGNGHFYRG 60
DB 29 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIDKSKQGGSGGDFDYE 88

QY 61 KASTDTMG 68
DB 89 KMANANKG 96

RESULT 43
US-09-219-019-17
; Sequence 17, Application US/09219019
; Patent No. 6268341
; GENERAL INFORMATION:
; APPLICANT: ROSENBERG, STEVEN
; TITLE OF INVENTION: EXPRESSION OF UROKINASE PLASMINOGEN ACTIVATOR
```

QY 1 SNELHQPNSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
Db 29 SNELHQPNSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 88  
QY 61 KASTDTMG 68  
Db 89 KMANANKG 96

RESULT 45  
PCT-US94-05669A-17  
; Sequence 17, Application PC/TUS9405669A  
; GENERAL INFORMATION:  
; APPLICANT: Chiron Corporation  
; TITLE OF INVENTION: Expression of Urokinase Plasminogen  
; TITLE OF INVENTION: Activator Inhibitors  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/05669A  
; FILING DATE: 19-MAY-1994  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Green, Grant D.  
; REGISTRATION NUMBER: 31,259  
; REFERENCE/DOCKET NUMBER: 0939.100  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 510-601-2706  
; TELEFAX: 510-655-3542  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 233 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US94-05669A-17

Query Match 37.8%; Score 300; DB 5; Length 233;  
Best Local Similarity 77.9%; Pred. No. 1.7e-22;  
Matches 53; Conservative 2; Mismatches 13; Indels 0; Gaps 0;  
QY 1 SNELHQPNSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
Db 29 SNELHQPNSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 98  
QY 61 KASTDTMG 68  
Db 89 KMANANKG 96

RESULT 46  
US-08-438-745-13  
; Sequence 13, Application US/08438745  
; Patent No. 6248715  
; GENERAL INFORMATION:  
; APPLICANT: Rosenber, Steven  
; APPLICANT: Stratton-Thomas, Jennifer  
; TITLE OF INVENTION: Expression of Urokinase Plasminogen  
; TITLE OF INVENTION: Activator Inhibitors  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/05669A  
; FILING DATE: 19-MAY-1994  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Green, Grant D.  
; REGISTRATION NUMBER: 31,259  
; REFERENCE/DOCKET NUMBER: 0939.100  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 510-601-2706  
; TELEFAX: 510-655-3542  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 233 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US94-05669A-17

Query Match 37.8%; Score 300; DB 5; Length 233;  
Best Local Similarity 77.9%; Pred. No. 1.7e-22;  
Matches 53; Conservative 2; Mismatches 13; Indels 0; Gaps 0;  
QY 1 SNELHQPNSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
Db 29 SNELHQPNSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 98  
QY 61 KASTDTMG 68  
Db 89 KMANANKG 96

RESULT 47  
US-09-219-019-13  
; Sequence 13, Application US/09219019  
; Patent No. 6268341  
; GENERAL INFORMATION:  
; APPLICANT: ROSENBERG, STEVEN  
; APPLICANT: STRATTON-THOMAS, JENNIFER R.  
; TITLE OF INVENTION: EXPRESSION OF UROKINASE PLASMINOGEN ACTIVATOR  
; TITLE OF INVENTION: INHIBITORS  
; FILE REFERENCE: 23533-0005  
; CURRENT APPLICATION NUMBER: US/09/219,019  
; CURRENT FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 08/438,263  
; PRIOR FILING DATE: 1995-05-10  
; PRIOR APPLICATION NUMBER: 08/280,288  
; PRIOR FILING DATE: 1994-07-26  
; PRIOR APPLICATION NUMBER: 08/070,153  
; PRIOR FILING DATE: 1993-06-01  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 235  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-219-019-13

Query Match 37.8%; Score 300; DB 3; Length 235;  
Best Local Similarity 77.9%; Pred. No. 1.7e-22;  
Matches 53; Conservative 2; Mismatches 13; Indels 0; Gaps 0;  
QY 1 SNELHQPNSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
Db 31 SNELHQPNSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 90  
QY 61 KASTDTMG 68  
Db 91 KMANANKG 98

RESULT 48  
US-09-219-019-13  
; Sequence 13, Application US/09219019  
; Patent No. 6268341  
; GENERAL INFORMATION:  
; APPLICANT: ROSENBERG, STEVEN  
; APPLICANT: STRATTON-THOMAS, JENNIFER R.  
; TITLE OF INVENTION: EXPRESSION OF UROKINASE PLASMINOGEN ACTIVATOR  
; TITLE OF INVENTION: INHIBITORS  
; FILE REFERENCE: 23533-0005  
; CURRENT APPLICATION NUMBER: US/09/219,019  
; CURRENT FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 08/438,263  
; PRIOR FILING DATE: 1995-05-10  
; PRIOR APPLICATION NUMBER: 08/280,288  
; PRIOR FILING DATE: 1994-07-26  
; PRIOR APPLICATION NUMBER: 08/070,153  
; PRIOR FILING DATE: 1993-06-01  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 235  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-219-019-13

ADDRESSEE: Chiron Corporation  
STREET: 4560 Horton Street  
CITY: Emeryville  
STATE: CA  
COUNTRY: USA  
ZIP: 94608  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/438,745  
FILING DATE: 10-MAY-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/070,153  
FILING DATE: 01-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Green, Grant D.  
REGISTRATION NUMBER: 31,259  
REFERENCE/DOCKET NUMBER: 0939.001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 510-601-2706  
TELEFAX: 510-655-3542  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 235 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-438-745-13

Query Match 37.8%; Score 300; DB 3; Length 235;  
Best Local Similarity 77.9%; Pred. No. 1.7e-22;  
Matches 53; Conservative 2; Mismatches 13; Indels 0; Gaps 0;  
QY 1 SNELHQPNSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
Db 31 SNELHQPNSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 90  
QY 61 KASTDTMG 68  
Db 91 KMANANKG 98

RESULT 49  
US-09-219-019-13  
; Sequence 13, Application US/09219019  
; Patent No. 6268341  
; GENERAL INFORMATION:  
; APPLICANT: ROSENBERG, STEVEN  
; APPLICANT: STRATTON-THOMAS, JENNIFER R.  
; TITLE OF INVENTION: EXPRESSION OF UROKINASE PLASMINOGEN ACTIVATOR  
; TITLE OF INVENTION: INHIBITORS  
; FILE REFERENCE: 23533-0005  
; CURRENT APPLICATION NUMBER: US/09/219,019  
; CURRENT FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 08/438,263  
; PRIOR FILING DATE: 1995-05-10  
; PRIOR APPLICATION NUMBER: 08/280,288  
; PRIOR FILING DATE: 1994-07-26  
; PRIOR APPLICATION NUMBER: 08/070,153  
; PRIOR FILING DATE: 1993-06-01  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 235  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-219-019-13

Query Match 37.8%; Score 300; DB 3; Length 235;  
Best Local Similarity 77.9%; Pred. No. 1.7e-22;  
Matches 53; Conservative 2; Mismatches 13; Indels 0; Gaps 0;  
QY 1 SNELHQPNSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
Db 31 SNELHQPNSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 90  
QY 61 KASTDTMG 68  
Db 91 KMANANKG 98

RESULT 50  
US-09-219-019-13  
; Sequence 13, Application US/09219019  
; Patent No. 6268341  
; GENERAL INFORMATION:  
; APPLICANT: ROSENBERG, STEVEN  
; APPLICANT: STRATTON-THOMAS, JENNIFER R.  
; TITLE OF INVENTION: EXPRESSION OF UROKINASE PLASMINOGEN ACTIVATOR  
; TITLE OF INVENTION: INHIBITORS  
; FILE REFERENCE: 23533-0005  
; CURRENT APPLICATION NUMBER: US/09/219,019  
; CURRENT FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 08/438,263  
; PRIOR FILING DATE: 1995-05-10  
; PRIOR APPLICATION NUMBER: 08/280,288  
; PRIOR FILING DATE: 1994-07-26  
; PRIOR APPLICATION NUMBER: 08/070,153  
; PRIOR FILING DATE: 1993-06-01  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 235  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-219-019-13

Query Match 37.8%; Score 300; DB 3; Length 235;  
Best Local Similarity 77.9%; Pred. No. 1.7e-22;  
Matches 53; Conservative 2; Mismatches 13; Indels 0; Gaps 0;  
QY 1 SNELHQPNSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
Db 31 SNELHQPNSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 90  
QY 61 KASTDTMG 68  
Db 91 KMANANKG 98

RESULT 51  
US-09-219-019-13  
; Sequence 13, Application US/09219019  
; Patent No. 6268341  
; GENERAL INFORMATION:  
; APPLICANT: ROSENBERG, STEVEN  
; APPLICANT: STRATTON-THOMAS, JENNIFER R.  
; TITLE OF INVENTION: EXPRESSION OF UROKINASE PLASMINOGEN ACTIVATOR  
; TITLE OF INVENTION: INHIBITORS  
; FILE REFERENCE: 23533-0005  
; CURRENT APPLICATION NUMBER: US/09/219,019  
; CURRENT FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: 08/438,263  
; PRIOR FILING DATE: 1995-05-10  
; PRIOR APPLICATION NUMBER: 08/280,288  
; PRIOR FILING DATE: 1994-07-26  
; PRIOR APPLICATION NUMBER: 08/070,153  
; PRIOR FILING DATE: 1993-06-01  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 235  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-219-019-13

Query Match 37.8%; Score 300; DB 3; Length 235;  
Best Local Similarity 77.9%; Pred. No. 1.7e-22;  
Matches 53; Conservative 2; Mismatches 13; Indels 0; Gaps 0;  
QY 1 SNELHQPNSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60  
Db 31 SNELHQPNSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 90  
QY 61 KASTDTMG 68  
Db 91 KMANANKG 98



Best Local Similarity 77.9%; Pred. No. 1.7e-22; Indels 0; Gaps 0;  
Matches 53; Conservative 2; Mismatches 13;

QY 1 SNELHVPNSCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCEIDKSKTCVEGNHGYRG 60  
DB 31 SNELHVPNSCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCEIDKSKTCVEGNHGYRG 90  
QY 61 KASTDTMG 68  
DB 91 KXANANKG 98

## RESULT 48

PCT-US94-05669A-13  
Sequence 13, Application PC/TUS9405669A

## GENERAL INFORMATION:

APPLICANT: Chiron Corporation

TITLE OF INVENTION: Expression of Urokinase Plasminogen

TITLE OF INVENTION: Activator Inhibitors

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Chiron Corporation

STREET: 4560 Horton Street

CITY: Emeryville

STATE: CA

COUNTRY: USA

ZIP: 94608

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/05669A

FILING DATE: 19-MAY-1994

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Green, Grant D.

REGISTRATION NUMBER: 31,259

REFERENCE/DOCKET NUMBER: 0939.100

TELECOMMUNICATION INFORMATION:

TELEPHONE: 510-601-2706

TELEFAX: 510-655-3542

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 235 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US94-05669A-13

Query Match 37.8%; Score 300; DB 5; Length 235;

Best Local Similarity 77.9%; Pred. No. 1.7e-22;

Matches 53; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY 1 SNELHVPNSCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCEIDKSKTCVEGNHGYRG 60  
DB 31 SNELHVPNSCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCEIDKSKTCVEGNHGYRG 90

QY 61 KASTDTMG 68

DB 91 KXANANKG 98

## RESULT 49

US-08-747-915-5

Sequence 5, Application US/08747915

Patent No. 5942492

GENERAL INFORMATION:

APPLICANT: Jones, Terence R.

APPLICANT: Haney, David N.

APPLICANT: Varga, Janos

TITLE OF INVENTION: CYCLIC PEPTIDES THAT BIND TO

TITLE OF INVENTION: UROKINASE-TYPE PLASMINOGEN ACTIVATOR RECEPTOR

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORRISON & FOERSTER

STREET: 2000 PENNSYLVANIA AVENUE, NW

CITY: WASHINGTON

STATE: DC

COUNTRY: USA

ZIP: 20006-1812

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/747,915

FILING DATE: 12-NOV-1996

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: MURASHIGE, KATE H.

REGISTRATION NUMBER: 29,959

REFERENCE/DOCKET NUMBER: 32904-20001.00

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500

TELEFAX: (202) 887-0763

TELEX: 90-4030 MRSNFOERSWSH

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 49 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:

NAME/KEY: Disulfide-bond

LOCATION: group(11..19, 13..31, 33..42)

US-08-747-915-5

Query Match

Best Local Similarity 36.9%; Score 293; DB 2; Length 49;

Mismatches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 SNELHVPNSCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCEIDKSKT 49  
DB 1 SNELHVPNSCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCEIDKSKT 49

## RESULT 50

US-09-285-783-5

Sequence 5, Application US/09285783

Patent No. 6514710

GENERAL INFORMATION:

APPLICANT: Jones, Terence R.

APPLICANT: Haney, David N.

APPLICANT: Varga, Janos

TITLE OF INVENTION: CYCLIC PEPTIDES THAT BIND TO

UROKINASE-TYPE PLASMINOGEN ACTIVATOR RECEPTOR

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: RADER, FISHMAN & GRAUER

STREET: 1233 20TH STREET NW, SUITE 501

CITY: WASHINGTON

STATE: DC

COUNTRY: USA

ZIP: 20036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/285,783

FILING DATE: 05-APR-1999

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:  
NAME: LIVNAT, SHMUEL  
REGISTRATION NUMBER: 33,949  
REFERENCE/DOCKET NUMBER: ANG-001/DIV (80144-0007)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 955-8787  
TELEFAX: (202) 955-3751  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 49 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: group(11..19, 13..31, 33..42)  
SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-09-285-783-5  
Query Match 36.9%; Score 293; DB 4; Length 49;  
Best Local Similarity 100.0%; Pred. No. 1.5e-22;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 SNELHQPNSNCDLNGGTCVSNKYPSNIHWNCNCPKPKFGGQHCIEDKSKT 49  
Db 1 SNELHQPNSNCDLNGGTCVSNKYPSNIHWNCNCPKPKFGGQHCIEDKSKT 49

Search completed: May 25, 2004, 15:00:01  
Job time : 10.3471 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2004, 14:47:10 ; Search time 15.9649 seconds

(without alignments)  
1662.947 Million cell updates/sec

Title: US-09-880-503-5

Perfect score: 1508

Sequence: 1 KPSPPEELKFCQGQKTLRP.....VSHFLPWRSHKENGIAL 276

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : PIR 78.\*

1: PIR1.\*  
2: PIR2.\*  
3: PIR3.\*  
4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1505	99.8	431	1 UKHU	u-plasminogen acti
2	1434.5	95.1	433	1 UKBAY	u-plasminogen acti
3	1255	83.2	442	1 UKPG	u-plasminogen acti
4	1188	78.8	433	1 JN0560	u-plasminogen acti
5	1119.5	74.2	432	1 S18932	u-plasminogen acti
6	1112.5	73.8	433	1 UKMS	u-plasminogen acti
7	688.5	45.7	434	1 A35005	u-plasminogen acti
8	583	38.7	582	1 UKHUT	t-plasminogen acti
9	570.5	37.8	559	1 A35029	t-plasminogen acti
10	569.5	37.8	559	1 A29941	t-plasminogen acti
11	547	36.3	431	2 JS0599	t-plasminogen acti
12	546	36.2	394	2 JS0600	t-plasminogen acti
13	545	36.2	477	2 JS0598	t-plasminogen acti
14	545	36.1	477	1 A34369	t-plasminogen acti
15	543	36.0	477	2 JS0597	t-plasminogen acti
16	495	32.8	615	1 A46H12	hepatocyte growth
17	493.5	32.7	655	1 A46688	coagulation factor
18	485.5	32.2	603	2 S28941	coagulation factor
19	453	30.0	593	2 S45281	brain-specific ser
20	425.5	28.2	761	2 JC5759	plasma hyaluronan-
21	425	28.2	558	2 JC5878	proctasin (EC 3.4.
22	417	27.7	580	1 JC4795	plasma kallikrein
23	408.5	27.1	343	1 A57014	plasma kallikrein
24	407	27.0	638	1 KQMSP	hepsin (EC 3.4.21
25	402	26.7	638	1 KQHUP	trypsin (EC 3.4.21
26	396.5	26.3	417	1 S00945	trypsin (EC 3.4.21
27	395.5	26.2	248	2 S55066	trypsin (EC 3.4.21
28	394.5	26.2	229	1 TRBQTR	trypsin (EC 3.4.21
29	392.5	26.0	460	2 B61545	plasmin (EC 3.4.21

30	390.5	25.9	247	2 S13813	trypsin (EC 3.4.21
31	390	25.9	263	2 A21195	chymotrypsin (EC 3
32	386.5	25.6	416	1 S33777	hepsin (EC 3.4.21
33	385	25.5	638	1 KQRTPL	plasma kallikrein
34	383.5	25.4	790	1 PLPG	plasmin (EC 3.4.21
35	382	25.3	263	1 KYRTB	chymotrypsin (EC 3
36	380	25.2	269	2 A26823	pancreatic elastase
37	379.5	25.2	304	2 S33496	chymotrypsin-like
38	378.5	25.1	264	2 I38136	chymotrypsin (EC 3
39	378	25.1	259	2 I38363	trypsin (EC 3.4.21
40	377	25.0	283	2 A31299	chymotrypsin (EC 3
41	376.5	25.0	247	1 A25852	trypsin (EC 3.4.21
42	375	24.9	812	1 PLMS	pancreatic elastase
43	374	24.8	271	1 ELRT2	trypsin (EC 3.4.21
44	373.5	24.8	247	1 TRDG	tissue kallikrein
45	373.5	24.8	261	2 A25606	7S nerve growth fa
46	372	24.7	261	1 NGMSG	pancreatic elastase
47	372	24.7	289	2 B26823	plasmin (EC 3.4.21
48	372	24.7	812	1 PLBO	chymotrypsin (EC 3
49	371	24.6	245	1 KYBOA	coagulation factor
50	370.5	24.6	625	1 KFHU1	plasmin (EC 3.4.21
51	370.5	24.6	810	1 PLHU	chymotrypsin (EC 3
52	370	24.5	245	1 KYBOB	trypsin (EC 3.4.21
53	369.5	24.5	246	1 TRRT1	trypsin (EC 3.4.21
54	368.5	24.4	238	2 S31779	trypsin (EC 3.4.21
55	367.5	24.4	231	1 TRPGR	trypsin (EC 3.4.21
56	367.5	24.4	243	2 A35871	serine proteinase
57	367.5	24.4	253	2 A53968	serine proteinase
58	366.5	24.3	237	2 S55378	trypsin (EC 3.4.21
59	366.5	24.3	248	2 S55067	trypsin (EC 3.4.21
60	366	24.3	274	2 JC4171	trypsin (EC 3.4.21
61	366	24.3	275	2 A32410	trypsin (EC 3.4.21
62	364.5	24.2	247	2 A27547	trypsin (EC 3.4.21
63	363.5	24.1	276	2 A38654	trypsin (EC 3.4.21
64	362.5	24.0	461	1 KKHU	trypsin (EC 3.4.21
65	362	24.0	855	2 JC7731	trypsin (EC 3.4.21
66	361.5	24.0	810	2 B30848	trypsin (EC 3.4.21
67	360	23.9	247	2 S05494	trypsin (EC 3.4.21
68	359	23.8	265	1 KQRP	tissue kallikrein
69	358.5	23.8	244	2 A44284	tissue kallikrein
70	358.5	23.8	246	1 TRRT2	trypsin (EC 3.4.21
71	358	23.7	261	2 S01971	tissue kallikrein
72	358	23.7	455	2 A61545	plasmin (EC 3.4.21
73	358	23.7	1524	2 T30337	polyprotein - Afri
74	358	23.7	4548	1 S00657	apoptotain(a) (EC
75	356.5	23.6	246	1 TRDGC	trypsin (EC 3.4.21

ALIGNMENTS

RESULT 1

UKHU

u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human

N/Alternate names: cellular plasminogen activator; urokinase; urokinase-type p.  
N/Contains: urokinase-type plasminogen activator chain A; urokinase-type plasminogen

in form

C/Species: Homo sapiens (man)

C/Date: 17-Dec-1992 #sequence revision 04-Dec-1986 #text change 15-Sep-2000

C/Accession: A00931; 152209; J70102; A37561; I38102; S65783; A37562; A37563; A

R/Riccio, A.; Grimaldi, G.; Verde, P.; Sebastio, G.; Boast, S.; Blas, F.

Nucleic Acids Res. 13, 2759-2771, 1985

A/Title: The human urokinase-plasminogen activator gene and its promoter.

A/Reference number: A00931; MUID:85215647; PMID:2987867

A/Accession: A00931

A/Molecule type: DNA

A/Residues: 1-431 <RIC>

A/Cross-references: GB:X02419; NID:937601; PIDN:CA26268.1; PID:gl834524

A/Note: the authors translated the codon ATG for residue 214 as Ile

R/Nagamine, Y.; Pearson, D.; Grattan, M.

Biochem. Biophys. Res. Commun. 132, 563-569, 1985

A/Title: Exon-intron boundary sliding in the generation of two mRNAs coding for

A/Reference number: 152209; MUID:86050639; PMID:3933505

A/Accession: I52209  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 145-161 <NAG1>  
 A/Cross-references: GB:X03027; NID:g340174; PIDN:AAA61257.1; PTD:g340175  
 R/Nagai, M.; Hiramatsu, R.; Kaneda, T.; Hayasuke, N.; Arimura, H.; Nishida, M.; Suyama, Gene 36, 183-188, 1995  
 A/Title: Molecular cloning of cDNA coding for human preprourokinase.  
 A/Reference number: J10102; MUID:86056954; PMID:2415429  
 A/Accession: J10102  
 A/Molecule type: mRNA  
 A/Residues: 1-213, 'I', 215-431 <NAG2>  
 A/Cross-references: GB:X03026; NID:g340155; PIDN:AA071138.1; PTD:g340158; GB:D00244; NID:Verde, P.; Stoppelli, M.P.; Galeffi, P.; Di Nocera, P.; Biasi, F. Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984  
 A/Title: Identification and primary sequence of an unspliced human urokinase poly(A) + RN  
 A/Reference number: A37561; MUID:84272706; PMID:6589620  
 A/Accession: A37561  
 A/Molecule type: mRNA  
 A/Residues: 66-431 <VER>  
 A/Cross-references: GB:D00244; NID:g220138  
 R/Jacobs, P.; Cravador, A.; Loriau, R.; Brockly, F.; Coliau, B.; Chuchana, P.; van Elsen, DNA 4, 139-146, 1985  
 A/Title: Molecular cloning, sequencing, and expression in Escherichia coli of human prepro  
 A/Reference number: I38102; MUID:85203359; PMID:3888571  
 A/Accession: I38102  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-150, 'W', 152-213, 'I', 215-385, 'C', 387-429, 'V', 431 <JAC>  
 A/Cross-references: EMBL:X02760; NID:g35297; PIDN:CAA26535.1; PID:g35298  
 R/Yoshimoto, M.; Ushiyama, Y.; Sakai, M.; Tamaki, S.; Hara, H.; Takahashi, K.; Sawasaki, Biochim. Biophys. Acta 1293, 83-89, 1996  
 A/Title: Characterization of single chain urokinase-type plasminogen activator with a no  
 A/Reference number: S65783; MUID:96186279; PMID:8652631  
 A/Accession: S65783  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 21-140, 'L', 142-213, 'I', 215-431 <YOS>  
 A/Cross-references: EMBL:D11143; NID:g1311467; PIDN:BA001919.1; PID:g1199928  
 R/Gunzler, W.A.; Steffens, G.J.; Otting, F.; Kim, S.M.A.; Frankus, E.; Flohe, L. Hoppe-Seyler's Z. Physiol. Chem. 363, 1155-1165, 1982  
 A/Title: The primary structure of high molecular mass urokinase from human urine.  
 A/Reference number: A37562; MUID:83055084; PMID:6754569  
 A/Accession: A37562  
 A/Molecule type: protein  
 A/Residues: 21-177 <GUN>  
 R/Schaller, J.; Nick, H.; Rickli, E.E.; Gillesen, D.; Lergier, W.; Studer, R.O. Eur. J. Biochem. 125, 251-257, 1982  
 A/Title: Human low-molecular-weight urinary urokinase. Partial characterization and prel  
 A/Reference number: A37563; MUID:83003608; PMID:6749491  
 A/Accession: A37563  
 A/Molecule type: protein  
 A/Residues: 156-176, 179-193, 'T', 195, 'T', 197-224 <SCH>  
 R/Steiffens, G.J.; Gunzler, W.A.; Otting, F.; Frankus, E.; Flohe, L. Hoppe-Seyler's Z. Physiol. Chem. 363, 1043-1058, 1982  
 A/Title: The complete amino acid sequence of low molecular mass urokinase from human uri  
 A/Reference number: A37564; MUID:83055099; PMID:6754572  
 A/Accession: A37564  
 A/Molecule type: protein  
 A/Residues: 158-410 <STE>  
 R/Kentzer, E.J.; Buko, A.; Menon, G.; Sarin, V.K. Biochem. Biophys. Res. Commun. 171, 401-406, 1990  
 A/Title: Carbohydrate composition and presence of a fucose-protein linkage in recombinan  
 A/Reference number: A35689; MUID:90365737; PMID:2393398  
 A/Accession: A35689  
 A/Molecule type: protein  
 A/Residues: 21-30, 'X', 32, 'X', 34-38, 'X', 40-43 <KEN>  
 R/Note: identification of a fucose and attempt to determine its attachment site  
 R/Rabbani, S.A.; Desjardins, J.; Bell, A.W.; Banville, D.; Mazar, A.; Henkin, J.; Goltzm Biochem. Biophys. Res. Commun. 173, 1058-1064, 1990  
 A/Title: An amino-terminal fragment of urokinase isolated from a prostate cancer cell li  
 A/Reference number: A36697; MUID:91097529; PMID:2125213  
 A/Accession: A36697

A/Molecule type: protein  
 A/Residues: 21-34 <RAB>  
 R/Li, X.; Bokman, A.M.; Llinas, M.; Smith, R.A.G.; Dobson, C.M. submitted to the Brookhaven Protein Data Bank, July 1993  
 A/Reference number: A51255; PDB:1KDU  
 A/Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR  
 R/Li, X.; Smith, R.A.G.; Dobson, C.M. Biochemistry 31, 9562-9571, 1992  
 A/Title: Sequential (1)H NMR assignments and secondary structure of the kringle  
 A/Reference number: A44375; MUID:93003110; PMID:1327118  
 A/Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR  
 R/Hansen, A.P.; Petros, A.M.; Meadows, R.P.; Nettlesheim, D.G.; Mazar, A.P.; Ol  
 submitted to the Brookhaven Protein Data Bank, January 1994  
 A/Reference number: A66822; PDB:1URK  
 A/Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR  
 R/Spaggan, G.S.; Phillips, C.; Nowak, U.K.; Ponting, C.P.; Saunders, D.; Dobs. submitted to the Brookhaven Protein Data Bank, July 1995  
 A/Reference number: A66058; PDB:1LMW  
 A/Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 169-175.  
 C/Comment: This enzyme is found in urine in a high molecular mass form, consis  
 C/Comment: Urokinase-type plasminogen activator proteolytically activates plas  
 C/Genetics:  
 A/Gene: GDB:PLAU  
 A/Cross-references: GDB:119497; OMIM:191840  
 A/Map position: 10Q24-10Q24  
 A/Introns: 19/3; 29/1; 123/2; 154/1; 227/2; 277/1; 324/1; 373/3  
 C/Function:  
 A/Description: proteolytically activates plasminogen  
 A/Pathway: fibrinolysis  
 C/Superfamily: urokinase-type plasminogen activator; EGF homology; kringle hom  
 C/Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serin  
 F:1-20/Domain: signal sequence #status predicted <SIG>  
 F:21-431/Product: urokinase-type plasminogen activator, single chain form #sta  
 F:21-177/Product: urokinase-type plasminogen activator chain A #status experim  
 F:21-62/Domain: EGF homology <EGF>  
 F:70-151/Domain: kringle homology <KRG>  
 F:156-177/Product: urokinase-type plasminogen activator chain A1 #status exper  
 F:179-431/Product: urokinase-type plasminogen activator chain B #status exper  
 F:179-419/Domain: trypsin homology <TRY>  
 F:31-39,33-51,53-62,70-151,91-133,122-146,168-299,209-225,217-288,313-382,345-  
 F:318/Binding site: carbohydrate (Thr) (covalent) #status predicted  
 F:178-179/Cleavage site: Lys-ile (plasmin) #status experimental  
 F:224,275,376/Active site: His, Asp, Ser #status experimental  
 F:222/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 Query Match 99.48; Score 1505; DB 1; Length 431;  
 Best Local Similarity 99.63; Pred. No. 4.2e-123;  
 Matches 275; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KPSSPPPEELKFCQCKTLRPRFKIIGGEFTTIENQPFMAIYRRHGGSVYVYCGGSLIS 60  
 DB 156 KPSSPPPEELKFCQCKTLRPRFKIIGGEFTTIENQPFMAIYRRHGGSVYVYCGGSLMS 215  
 QY 61 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGEWKEVENLILHKDYSADTLAHHND 120  
 DB 216 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGEWKEVENLILHKDYSADTLAHHND 275  
 QY 121 IALLKIRSKRGCAQPSRTTQTICLPMSYNDPFGTSCETITGFKENSTDTLYPEQLKMT 180  
 DB 276 IALLKIRSKRGCAQPSRTTQTICLPMSYNDPFGTSCETITGFKENSTDTLYPEQLKMT 335  
 QY 181 VYKLISHRECOOPHYGSEVTTVMKCAADPQWTDSCQDSCGFLVCSLOGRMTLTGIVS 240  
 DB 336 VYKLISHRECOOPHYGSEVTTVMKCAADPQWTDSCQDSCGFLVCSLOGRMTLTGIVS 395  
 QY 241 WGRCCALKDKEGVYTRVSHFLPWRSHTKENGIAL 276  
 DB 396 WGRCCALKDKEGVYTRVSHFLPWRSHTKENGIAL 431  
 RESULT 2  
 UKBAY  
 u-plasminogen activator (EC 3.4.21.73) precursor - yellow baboon

C:Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)  
 C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 18-Jun-1999  
 C:Accession: S14687; S08651  
 R:Au, Y.P.T.; Wang, T.W.; Clowes, A.W.  
 Nucleic Acids Res. 18, 3411, 1990  
 A:Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminogen  
 A:Reference number: S14687; MUID:90287734; PMID:2113276  
 A:Accession: S14687  
 A:Molecule type: mRNA  
 A:Residues: 1-433 <AUY>  
 A:Cross-references: EMBL:X51935; NID:G38130; PIDN:CAA36200.1; PID:G38131  
 A:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try  
 C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase  
 F:1-20/Domain: signal sequence #status predicted <SIG>  
 F:21-176/Product: plasminogen activator chain A #status predicted <ACH>  
 F:30-61/Domain: EGF homology <EGF>  
 F:69-150/Domain: kringle homology <KRG>  
 F:178-433/Product: plasminogen activator chain B #status predicted <BCH>  
 F:178-421/Domain: trypsin homology <TRY>  
 F:167-298, 208-224, 216-287, 315-384, 374-402/Disulfide bonds: #status predicted  
 F:223, 274, 378/Active site: His, Asp, Ser #status predicted  
 F:324/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 95.1%; Score 1434.5; DB 1; Length 433;  
 Best Local Similarity 92.8%; Pred. No. 5.8e-117; Mismatches 4; Indels 3; Gaps 1;  
 Matches 259; Conservative 13;

QY 1 KPSSPEELKFCQGKTLRPFKIIGGFTTIENQPFPAAYRRHGGSVTVVCGSLIS 60  
 DB 155 KPSSPEELKFCQGKTLRPFKIIGGFTTIENQPFPAAYRRHGGSVTVVCGSLIS 214  
 QY 61 PCWISATGCHFDYPKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 120  
 DB 215 PCWVSATGCHFDYPKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 274  
 QY 121 IALLKIRSEGRCAQPSRTIOTICLPSMYNDPQ---FGTSCITGFGKNSDLYPRL 177  
 DB 275 IALLKIRSEGRCAQPSRTIOTICLPSMYNDPNDPFGTSCITGFGKNSDLYPRL 334  
 QY 178 KMTVVKLVSHRCQPHYGVSEVTTKMLCAADPOWKTDCQDGGPLVCSLQGRMTLTG 237  
 DB 335 KMTVVKLVSHRCQPHYGVSEVTTKMLCAADPOWKTDCQDGGPLVCSLQGRMTLTG 394  
 QY 238 IVSWRGCGALKDKPGVYTVRSHFLPWIRSHTKENGLAL 276  
 DB 395 IVSWRGCGALKDKPGVYTVRSHFLPWIRSHTKENGLAL 433

RESULT 3  
 UKPG  
 u-plasminogen activator (EC 3.4.21.73) precursor - pig  
 N:Alternate names: uPA  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 04-Dec-1986 #sequence\_revision 17-Mar-1987 #text\_change 07-Aug-1998  
 C:Accession: A00932  
 R:Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, E.  
 Nucleic Acids Res. 12, 9525-9541, 1984  
 A:Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.  
 A:Reference number: A00932; MUID:85067954; PMID:6096832  
 A:Accession: A00932  
 A:Molecule type: DNA  
 A:Residues: 1-240, 'H', 242-442 <NAG1>  
 A:Experimental source: kidney cell line LLC-PK1  
 R:Nagamine, Y.  
 submitted to the Protein Sequence Database, December 1986  
 A:Reference number: A37566  
 A:Contents: annotation; correction to residue 241  
 A:Genetics:  
 A:Introns: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3  
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try  
 C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase  
 F:1-20/Domain: signal sequence #status predicted <SIG>  
 F:21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>

F:33-64/Domain: EGF homology <EGF>  
 F:72-153/Domain: kringle homology <KRG>  
 F:190-442/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>  
 F:190-430/Domain: trypsin homology <TRY>  
 F:152/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:178-310, 220-236, 228-299, 324-393, 356-372, 383-411/Disulfide bonds: #status predicted  
 F:235, 286, 387/Active site: His, Asp, Ser #status predicted

Query Match 83.2%; Score 1255; DB 1; Length 442;  
 Best Local Similarity 82.1%; Pred. No. 2.4e-101; Mismatches 25; Indels 0; Gaps 0  
 Matches 225; Conservative 23;

QY 2 PSSPPEELKFCQGKTLRPFKIIGGFTTIENQPFPAAYRRHGGSVTVVCGSLIS 61  
 DB 168 PSSPPEELKFCQGKTLRPFKIIGGFTTIENQPFPAAYRRHGGSVTVVCGSLIS 227  
 QY 62 CWVISATGCHFDYPKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 121  
 DB 228 CWVSATGCHFDYPKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 287  
 QY 122 ALLKIRSEGRCAQPSRTIOTICLPSMYNDPQGTSCITGFGKNSDLYPRLQKMTV 181  
 DB 288 ALLKIRTDKCCQAPSRSIOTICLPVNGDAHFGASCEIVGFKEDPSDLYPRLQKMTV 347  
 QY 182 VKLISHRECCQPHYGVSEVTTKMLCAADPOWKTDCQDGGPLVCSLQGRMTLTGIVSW 241  
 DB 348 VKLISHRECCQPHYGVSEVTTKMLCAADPOWKTDCQDGGPLVCSLQGRMTLTGIVSW 407  
 QY 242 GRGKALKDKPGVYTVRSHFLPWIRSHTKENGLA 275  
 DB 408 GRGKALKDKPGVYTVRSHFLPWIRSHTKENGLA 441

RESULT 4  
 JN0560  
 u-plasminogen activator (EC 3.4.21.73) precursor - bovine  
 N:Alternate names: uPA  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 16-Jul-1999  
 C:Accession: JN0560  
 R:Kraetzschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schleuning, W.D.  
 Gene 125, 177-183, 1993  
 A:Title: Bovine urokinase-type plasminogen activator and its receptor: cloning  
 A:Reference number: JN0560; MUID:93216119; PMID:8385052  
 A:Accession: JN0560  
 A:Molecule type: mRNA  
 A:Residues: 1-433 <KRA>  
 A:Cross-references: GB:L03546; NID:G163800; PIDN:AAA51419.1; PID:G163801  
 C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle hom  
 C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase  
 F:1-20/Domain: signal sequence #status predicted <SIG>  
 F:21-179/Product: plasminogen activator chain A #status predicted <MA1>  
 F:21-179/Product: urokinase-type plasminogen activator chain B #status predicted  
 F:33-64/Domain: EGF homology <EGF>  
 F:72-153/Domain: kringle homology <KRG>  
 F:181-433/Product: plasminogen activator chain B #status predicted <MA2>  
 F:181-421/Domain: trypsin homology <TRY>  
 F:170-301, 211-227, 219-290, 315-384, 347-363, 374-402/Disulfide bonds: #status pre  
 F:226, 277, 378/Active site: His, Asp, Ser #status predicted

Query Match 78.8%; Score 1188; DB 1; Length 433;  
 Best Local Similarity 76.4%; Pred. No. 1.6e-95; Mismatches 32; Indels 0; Gaps 0  
 Matches 210; Conservative 32;

QY 2 PSSPPEELKFCQGKTLRPFKIIGGFTTIENQPFPAAYRRHGGSVTVVCGSLIS 61  
 DB 159 PSSPPEELKFCQGKTLRPFKIIGGFTTIENQPFPAAYRRHGGSVTVVCGSLIS 218  
 QY 62 CWVISATGCHFDYPKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 121  
 DB 219 CWVSATGCHFDYPKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 278  
 QY 122 ALLKIRSEGRCAQPSRTIOTICLPSMYNDPQGTSCITGFGKNSDLYPRLQKMTV 181

Db 279 ALLKIRTSRQCAQPSRSITQICLPPEHEDAHSTRCEITGFKENPDSYRYSDELKMTF 338  
QY 182 VKLISHRECCQPHYGVSEVTTKMLCAADPOKWTDCSQSGGGLVCSLOGRWTLTGIVSW 241  
Db 339 VSLVSHVCCQPHYGVSEVTTKMLCAADPOKWTDCSQSGGGLVCSLOGRWTLTGIVSW 398  
QY 242 GRGCAKDKPGVYTRVSHFLPWIRSHTKKEENGLA 276  
Db 399 GRDCAMKPKGVYTRVSHFLPWIRSHTKKEENGLA 433

## RESULT 5

S18932  
u-plasminogen activator (EC 3.4.21.73) precursor - rat  
A:Alternate names: plasminogen activator, urokinase-type; urinary plasminogen activator  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 18-Oct-1989 #sequence revision 10-Feb-1995 #text\_change 18-Jun-1999  
C:Accession: S24604; I60186; I53472; S18932  
R:Rabbani, S.A.  
submitted to the EMBL Data Library, April 1992  
A:Reference number: S24604  
A:Accession: S24604  
A:Molecule type: mRNA  
A:Residues: 1-15,'H',17-23,'G',25-331,'N',333-432 <RAB>  
A:Cross-references: EMBL:X65651; NID:G57456; PIDN:CAA46601.1; PID:G57457  
A:Experimental source: tissue kidney  
R:Henderson, B.R.; Tansley, W.P.; Phillips, S.M.; Ramshaw, I.A.; Kefford, R.P.  
Cancer Res. 52, 2489-2496, 1992  
A:Title: Transcriptional and posttranscriptional activation of urokinase plasminogen act  
A:Reference number: I60186; MUID:92233409; PMID:1568219  
A:Accession: I60186  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-432 <RES>  
A:Cross-references: EMBL:X63434; NID:G57465; PIDN:CAA45028.1; PID:G57466  
A:Experimental source: strain Fischer 344; tissue mammary  
R:Ragno, P.; Cassano, S.; Degen, J.; Kessler, C.; Blasi, F.; Rossi, G.  
FEBS Lett. 306, 193-198, 1992  
A:Title: The receptor for the plasminogen activator of urokinase type is up-regulated in  
A:Reference number: I53472; MUID:92339549; PMID:1321734  
A:Accession: I53472  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 31-62 <RE2>  
A:Cross-references: EMBL:X66907; NID:G396200; PIDN:CAA47356.1; PID:G938279  
A:Gene: uPA  
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try  
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-177/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>  
F:31-62/Domain: EGF homology <EGF>  
F:70-151/Domain: kringle homology <KRG>  
F:179-432/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>  
F:179-420/Domain: trypsin homology <TRY>  
F:168-300,210-226,215-289,314-383,346-362,373-401/Disulfide bonds: #status predicted  
F:225,276,377/Active site: His, Asp, Ser #status predicted

Query Match 74.2%; Score 1119.5; DB 1; Length 432;  
Best Local Similarity 71.4%; Pred. No. 1.4e-89;  
Matches 197; Conservative 39; Mismatches 39; Indels 1; Gaps 1;  
QY 1 KPSSPPEELKFCQGGKTLPRFKIIGBFTTIENQWFAAIYRRHGGG-VTVVCGGSLI 59  
Db 156 KPSSVVDQGGFCQGGKALPRFKIVGFEFTVENQWFAAIYLRKNGGSPSPFKCGGSLI 215  
QY 60 SPCWVISAHCFTDYPKEDYIYVLGRSLNSNTQGMKFEVENILHKDYSADTLAHNN 119  
Db 216 SPCWVASATHCFTVQPKKEYYVVLGSKXNSNPNCEMKFEVEQLILHEDFSDETLAFN 275  
QY 120 DIALLKIRSEKGRCAQPSRTIQTICLPSPMYNDPQFTSCBITGFKENSTDYLYPEQLKM 179

Db 276 DIALLKIRTSRQCAQPSRSITQICLPPEFDAPFGSDCEITGQBSADTYFVPKDLKM 335  
QY 180 TVVKLIHSHRECCQPHYGVSEVTTKMLCAADPOKWTDCSQSGGGLVCSLOGRWTLTGIV 239  
Db 336 SVVKLIHSHRECCQPHYGVSEVTTKMLCAADPOKWTDCSQSGGGLVCSLOGRWTLTGIV 395  
QY 240 SWRGCAKDKPGVYTRVSHFLPWIRSHTKKEENGLA 275  
Db 396 SWRGCAKDKPGVYTRVSHFLPWIRSHTKKEENGLA 431

## RESULT 6

UROMS  
u-plasminogen activator (EC 3.4.21.73) precursor - mouse  
A:Species: Mus musculus (house mouse)  
C>Date: 30-Sep-1987 #sequence revision 30-Sep-1987 #text\_change 18-Jun-1999  
C:Accession: A29420; A24615  
R:Degen, S.J.F.; Heckel, J.L.; Reich, E.; Degen, J.L.  
Biochemistry 26, 8270-8279, 1987  
A:Title: The murine urokinase-type plasminogen activator gene.  
A:Reference number: A29420; MUID:88163489; PMID:2831940  
A:Accession: A29420  
A:Molecule type: DNA  
A:Residues: 1-433 <DEG>  
A:Cross-references: GB:X17922; NID:G202296; PIDN:AAA40539.1; PID:G202297  
R:Belin, D.; Vassalli, J.D.; Combepine, C.; Godeau, F.; Nagamine, Y.; Reich, E.  
Eur. J. Biochem. 148, 225-232, 1985  
A:Title: Cloning, nucleotide sequencing and expression of cDNAs encoding mouse  
A:Reference number: A24615; MUID:85179474; PMID:2985383  
A:Accession: A24615  
A:Molecule type: mRNA  
A:Residues: 1-433 <BEL>  
A:Cross-references: GB:X02389; NID:G55127; PIDN:CAA26231.1; PID:G55128  
C:Genetics:  
A:Insertions: 19/3; 30/1; 66/1; 124/2; 155/1; 229/2; 279/1; 326/1; 375/3  
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle hom  
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase  
F:1-20/Domain: signal sequence #status predicted <SIG>  
F:21-178/Product: urokinase-type plasminogen activator chain A #status predict  
F:32-63/Domain: EGF homology <EGF>  
F:71-152/Domain: kringle homology <KRG>  
F:180-433/Product: urokinase-type plasminogen activator chain B #status predic  
F:180-421/Domain: trypsin homology <TRY>  
F:169-301,211-227,219-290,315-384,347-363,374-402/Disulfide bonds: #status pre  
F:226,277,378/Active site: His, Asp, Ser #status predicted

Query Match 73.8%; Score 1112.5; DB 1; Length 433;  
Best Local Similarity 71.0%; Pred. No. 5.7e-89;  
Matches 196; Conservative 39; Mismatches 40; Indels 1; Gaps 1  
QY 1 KPSSPPEELKFCQGGKTLPRFKIIGBFTTIENQWFAAIYRRHGGG-VTVVCGGSLI 59  
Db 157 KPSSSVVDQGGFCQGGKALPRFKIVGFEFTVENQWFAAIYLRKNGGSPSPFKCGGSLI 216  
QY 60 SPCWVISAHCFTDYPKEDYIYVLGRSLNSNTQGMKFEVENILHKDYSADTLAHNN 119  
Db 217 SPCWVASAAHCFTQPKENYVVLGSKXESSNPNCEMKFEVEQLILHYYRSDSLAYHN 276  
QY 120 DIALLKIRSEKGRCAQPSRTIQTICLPSPMYNDPQFTSCBITGFKENSTDYLYPEQLKM 179  
Db 277 DIALLKIRTSRQCAQPSRSITQICLPPEFDAPFGSDCEITGFKESBSDLYLPKMLKM 336  
QY 180 TVVKLIHSHRECCQPHYGVSEVTTKMLCAADPOKWTDCSQSGGGLVCSLOGRWTLTGIV 239  
Db 337 SVVKLIHSHRECCQPHYGVSEVTTKMLCAADPOKWTDCSQSGGGLVCSLOGRWTLTGIV 396  
QY 240 SWRGCAKDKPGVYTRVSHFLPWIRSHTKKEENGLA 275  
Db 397 SWRGCAKDKPGVYTRVSHFLPWIRSHTKKEENGLA 432

## RESULT 7

A35005

u-plasminogen activator (EC 3.4.21.73) precursor - chicken  
N;Alternate names: uPA  
C;Species: Gallus gallus (chicken)  
C;Date: 20-Jul-1990 #sequence\_revision 20-Jul-1990 #text\_change 16-Jul-1999  
C;Accession: A35005  
R;Residue: N.D.; Kessler, C.A.; Bell, S.M.; Degen, J.L.  
J. Biol. Chem. 265, 1339-1344, 1990  
A;Title: The chicken urokinase-type plasminogen activator gene.  
A;Reference number: A35005; MUID:90110185; PMID:2295632  
A;Accession: A35005  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-434 <ES>  
A;Cross-references: GB:J05187; NID:9212858; PIDN:AAA49131.1; PID:9212859  
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try  
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;22-171/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>  
F;40-71/Domain: EGF homology <EGF>  
F;79-158/Domain: kringle homology <KRG>  
F;173-428/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>  
F;173-416/Domain: trypsin homology <TRY>  
F;162-296,202-218,210-285,310-379,342-358,369-397/Disulfide bonds: #status predicted  
F;217,272,373/Active site: His, Asp, Ser #status predicted  
  
Query Match 45.7%; Score 688.5; DB 1; Length 434;  
Best Local Similarity 47.4%; Pred. NO. 4.2e-52;  
Matches 128; Conservative 48; Mismatches 89; Indels 5; Gaps 2;  
  
QY 2 PSSPPEELKPCQCKTLRFRFKIIGGEFTTIENQPFALYRRHGGSVYVCGGSLISP 61  
Db 151 PCSTIEKERTCGRSFKYKLVGGSGAEVETQPIAGIFQNM-GTDQFLCGSLIDP 209  
  
QY 62 CWVISATHCID----YPKEDIVVLGRSLNSNTQGMKEFEVENLIUHKOYSADTLAH 117  
Db 210 CWLTLAAHCYFNPTKQPNKSVYKFLGKSLTNDHEQVPMVDIISHPFDTHGTGN 269  
  
QY 118 HNDIALKIRKSGRCACQPSRTTQICLPVMNDPQFGTSCITGKGNSTDYLYPEOL 177  
Db 270 DNDIALIRITASQCAVESNYVTVCLPEKNLNDNTWELAGYKGNKSDIYYAQL 329  
  
QY 178 KMTVKLIHRECCQPHYGVSEVTTKMLCAADPQWKTDCQDGGGFLVCSLQGRMTLTG 237  
Db 330 MSATVNLISQDDCKNYDSTRVTDNNVCAGDPLWETDCKGDSGPMVCEHNGRMTLYG 389  
  
QY 238 IVSWGRCALKKDPGVYTVRVSHLPWRSH 267  
Db 390 IVSWGRCALKKDPGVYTVRVRLNWDN 419  
  
RESULT 8  
UKHUT  
t-plasminogen activator (EC 3.4.21.68) precursor [validated] - human  
N;Alternate names: t-PA; tissue plasminogen activator  
C;Species: Homo sapiens (man)  
C;Date: 14-Nov-1983 #sequence\_revision 14-Nov-1983 #text\_change 08-Dec-2000  
A;Accession: A94004; A23529; J0562; A93293; S02125; A91343; A93951; A91322; A54645; I60  
R;NY, T.; Elgh, F.; Lund, B.  
Proc. Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984  
A;Title: The structure of the human tissue-type plasminogen activator gene: correlation  
A;Reference number: A94004; MUID:84298137; PMID:6089139  
A;Accession: A94004  
A;Molecule type: DNA  
A;Residues: 1-562 <NYT>  
A;Cross-references: GB:I00141  
A;Note: the codon given for residue 93 (ACC) is inconsistent with the authors' translation  
R;Frieze, Degen, S.J.; Rajput, B.; Reich, E.  
J. Biol. Chem. 261, 6972-6985, 1986  
A;Title: The human tissue plasminogen activator gene.  
A;Reference number: A23529; MUID:86196143; PMID:3009482  
A;Accession: A23529  
A;Molecule type: DNA  
A;Residues: 1-562 <DEG>

A;Cross-references: GB:K03021; NID:9339817; PIDN:AAA98809.1; PID:9339818  
R;Itagaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuda, S.; Higashio, K.  
Agric. Biol. Chem. 55, 1225-1232, 1991  
A;Title: Purification and characterization of tissue plasminogen activator secreted by  
A;Reference number: J0562; MUID:91291340; PMID:1368681  
A;Accession: J0562  
A;Molecule type: mRNA  
A;Residues: 31-562 <ITA>  
A;Cross-references: DDBJ:D01096; NID:9220128; PIDN:BA000881.1; PID:9441174  
A;Experimental source: embryonic lung fibroblast IMR-90 cells  
A;Note: part of this sequence, including the amino end of the mature protein, is confirmed  
R;Pennica, D.; Holmes, W.E.; Kohr, W.J.; Harkins, R.N.; Vehar, G.A.; Ward, C.A.  
Nature 301, 214-221, 1983  
A;Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Escherichia  
A;Reference number: A93293; MUID:83115262; PMID:6337343  
A;Accession: A93293  
A;Molecule type: mRNA  
A;Residues: 1-562 <PEN>  
A;Cross-references: GB:L00141  
A;Experimental source: melanoma cells  
R;Sasaki, H.; Saito, Y.; Hayashi, M.; Otsuka, K.; Niwa, M.  
Nucleic Acids Res. 16, 5695, 1988  
A;Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA from human  
A;Reference number: S02125; MUID:89285279; PMID:3133640  
A;Accession: S02125  
A;Status: translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-562 <SAS>  
A;Cross-references: EMBL:X07393; NID:937243; PIDN:CAA30302.1; PID:937244  
A;Experimental source: fetal lung cells  
R;Kagitani, H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Saito, A.; Bando, H.; Oka, K.;  
FEBS Lett. 189, 145-149, 1985  
A;Title: Expression in Escherichia coli of finger-domain lacking tissue-type plasminogen  
A;Reference number: A91343; MUID:85285620; PMID:3896853  
A;Accession: A91343  
A;Molecule type: mRNA  
A;Residues: 1-38, 'G', 86-433, 'E', 435-562 <KAG>  
A;Experimental source: Detroit 562 cells; ATCC 138  
R;Edlund, T.; Ny, T.; Ranby, M.; Heden, L.O.; Palm, G.; Holmgren, E.; Josephson, S.  
Proc. Natl. Acad. Sci. U.S.A. 80, 349-352, 1983  
A;Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen activator  
A;Reference number: A93951; MUID:83169656; PMID:6572897  
A;Accession: A93951  
A;Molecule type: mRNA  
A;Residues: 251-358 <EDL>  
A;Experimental source: melanoma cells  
R;Pohl, G.; Kallstrom, M.; Bergsdorf, N.; Wallen, P.; Jornvall, H.  
Biochemistry 23, 3701-3707, 1984  
A;Title: Tissue plasminogen activator: peptide analyses confirm an indirectly derived  
differences.  
A;Reference number: A90488; MUID:85000469; PMID:6433976  
A;Contents: annotation; melanoma cells, partial sequence of residues 36-562, active ar  
R;Pohl, G.; Kaplan, L.; Einarsson, M.; Wallen, P.; Jornvall, H.  
FEBS Lett. 168, 29-32, 1984  
A;Title: Differences between uterine and melanoma forms of tissue plasminogen activator  
A;Reference number: A91322; MUID:84158956; PMID:6538514  
A;Accession: A91322  
A;Molecule type: protein  
A;Residues: 33-45,311-320 <POH>  
A;Experimental source: uterus  
A;Note: in the uterus, cleavage of the activation peptide may also occur after  
R;van Zonneveld, A.J.; Veerman, H.; Pannekoek, H.  
J. Biol. Chem. 261, 14214-14218, 1986  
A;Reference number: A37567; MUID:87033611; PMID:3021732  
A;Contents: annotation; fibrin binding site  
R;Verheijen, J.H.; Caspers, M.F.M.; Chang, G.T.G.; de Munk, G.A.W.; Pouwels, P.; Eng  
EMBO J. 5, 3525-3530, 1986  
A;Title: Involvement of finger domain and kringle 2 domain of tissue-type plasminogen  
A;Reference number: A37568; MUID:87161761; PMID:3030730  
A;Contents: annotation; fibrin binding site  
R;Dodd, I.; Nunn, B.; Robinson, J.H.  
Thromb. Haemost. 59, 523-528, 1988  
A;Title: Isolation, identification and pharmacokinetic properties of human tissue-type



A:Reference number: A60902; MUID:89044681; PMID:3142086  
A:Contents: annotation; novel forms of expressed recombinant t-PA  
R;Harris, T.J.R.; Patel, T.; Warson, F.A.O.; Little, S.; Entage, J.S.; Opdenakker, G.;  
Mol. Biol. Med. 3, 279-292, 1986  
A:Title: Cloning of cDNA coding for human tissue-type plasminogen activator and its expression  
A:Reference number: A54645; MUID:86284200; PMID:3060401  
A:Accession: A54645  
A:Molecule type: mRNA  
A:Residues: 1-562 <HAR>  
A:Cross-references: GB:M15518; NID:G190031; PIDN:AAA60111.1; PID:G190032  
A:Note: parts of this sequence were confirmed by peptide sequencing  
R;Reddy, V.B.; Garramone, A.J.; Sasak, H.; Wei, C.  
DNA 6, 461-472, 1987  
A:Title: Expression of human uterine tissue-type plasminogen activator in mouse cells using  
A:Reference number: I60110; MUID:18054470; PMID:2824147  
A:Accession: I60110  
A:Status: translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-562 <RES>  
A:Cross-references: GB:M18182; NID:G340176; PIDN:AAA36800.1; PID:G340177  
R;Fisher, R.; Waller, E.K.; Grossi, G.; Thompson, D.; Tizard, R.; Schleuning, W.D.  
J. Biol. Chem. 260, 11223-11230, 1985  
A:Title: Isolation and characterization of the human tissue-type plasminogen activator gene  
A:Reference number: I55232; MUID:85289338; PMID:3161893  
A:Accession: I55232  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-36 <RES>  
A:Cross-references: GB:M11890; NID:G339837; PIDN:AAA61213.1; PID:G339839  
A:Comment: Cleavage by plasmin or trypsin produces two chains held together by a single  
C:Comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-Val bond. It  
C:Comment: t-PA binds chain A of fibrin by kringle 2 and the fibronectin type I repeat.  
C:Genetics:  
A:Gene: GDB:PLAT  
A:Cross-references: GDB:119496; OMIM:173370  
A:Map position: 9c12-8p12  
A:Exons: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2; 297/1; 362/2; 408/1; 455/3; 510  
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom  
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:24-32/Domain: propeptide #status predicted <PRO>  
F:33-562/Product: t-plasminogen activator #status experimental <MAT>  
F:33-310/Product: t-plasminogen activator chain A #status experimental <ACH>  
F:41-78/Domain: fibronectin type I repeat homology <RF1>  
F:86-119/Domain: EGF homology <EGF>  
F:127-208/Domain: kringle homology <KR1>  
F:215-296/Domain: kringle homology <KR2>  
F:311-562/Product: t-plasminogen activator chain B #status experimental <BCH>  
F:41-71,69-78,86-97,91-108,110-119,127-208,148-190,179-203,215-296,236-278,267-291,299-4  
F:152,483/Binding site: carboxydrate (Asn) [covalent] #status experimental  
F:219/Binding site: carboxydrate (Asn) [covalent] (partial) #status experimental  
F:310-311/Cleavage site: Arg-ile (plasmin, trypsin) #status experimental  
F:357,406/Active site: His, Asp #status predicted  
F:513/Active site: Ser #status experimental

Query Match 38.78; Score 583; DB 1; Length 562;  
Best Local Similarity 44.98; Pred. No. 8.5e-43;  
Matches 11; Conservative 39; Mismatches 97; Indels 10; Gaps 5;  
Qy 13 CG-QKTLR-PRFKIIGGFTTIENQPFPAIY-RRHGGSVTVYVCGGLISPCWVISATHC 70  
Db 299 CGLRQYQPOQFRKIGGLFADIASHPQAAIFVXRRSPGFRFLCGGILISSCWILSAHC 358  
Qy 71 FIDYPKKEDYIVYGRSLNSNTQGENKFEVENILHKDYSADTLAHNDIALKIRSK 130  
Db 359 FQRFPFHLLTVILGRYRVVPGEEQKFEVEKYVHKFEFDDT--YNDIALQLKSDS 416  
Qy 131 GRCAQPSRTIQTICLPSMYNDPFGTSCEITGFGKENSIDYLYPEQLKMTVVKLISHREC 190  
Db 417 SRCAQESSVVRTCLPADQLPDWTECELSGYGKHEALSPFFSRLEKHAHVRLYESSRC 476  
Qy 191 QQHYGSEVTTMCLAAD-----PWKT-DSQCQSGGGLVCSLQGRMTLTGIVSWGRG 244

Db 477 TSQHLLNRTVTDNMLCAGDTRSGPQANLHDCQSGGGLVCSLQGRMTLTGIVSWGLG 536  
Qy 245 CALKXKPGVTVTRVSHFLPWRSHTK 269  
Db 537 CGQKXQVGVTVKTVNYLDWDRDNR 561  
RESULT 9  
A35029  
t-plasminogen activator (EC 3.4.21.68) precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A35029; A31597  
R;Feng, P.; Ohlsson, M.; Ny, T.  
J. Biol. Chem. 265, 2022-2027, 1990  
A:Title: The structure of the TATA-less rat tissue-type plasminogen activator gene  
A:Reference number: A35029; MUID:90130448; PMID:2105315  
A:Accession: A35029  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-559 <FEN>  
A:Cross-references: GB:M31197; NID:G207429; PIDN:AAA42261.1; PID:G207431; GB:J 126  
R;NY, T.; Lehardsson, G.; Hsueh, A.J.W.  
DNA 7, 671-677, 1988  
A:Title: Cloning and characterization of a cDNA for rat tissue-type plasminogen activator  
A:Reference number: A31597; MUID:89170114; PMID:3148445  
A:Accession: A31597  
A:Molecule type: mRNA  
A:Residues: 1-379; 'K', 381-559 <NYT>  
A:Cross-references: GB:M23697; NID:G30159; PIDN:AAA41812.1; PID:G30160  
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat  
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F:1-17/Domain: signal sequence #status predicted <SIG>  
F:18-29/Domain: propeptide #status predicted <PRO>  
F:30-553/Product: t-plasminogen activator #status predicted <MAT>  
F:30-308/Product: t-plasminogen activator chain A #status predicted <ACH>  
F:38-75/Domain: fibronectin type I repeat homology <RF1>  
F:83-116/Domain: EGF homology <EGF>  
F:124-205/Domain: kringle homology <KR1>  
F:213-294/Domain: kringle homology <KR2>  
F:309-553/Product: t-plasminogen activator chain B #status predicted <BCH>  
F:309-553/Domain: trypsin homology <TRY>  
F:348-68,66-75,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265  
F:149,481/Binding site: carboxydrate (Asn) [covalent] #status predicted  
F:308-309/Cleavage site: Arg-ile (plasmin, trypsin) #status predicted  
F:355,404,510/Active site: His, Asp, Ser #status predicted  
Query Match 37.88; Score 570.5; DB 1; Length 559;  
Best Local Similarity 42.48; Pred. No. 1e-41;  
Matches 114; Conservative 47; Mismatches 91; Indels 17; Gaps 6  
Qy 13 CGQKTLR-PRFKIIGGFTTIENQPFPAIY-RRHGGSVTVYVCGGLISPCWVISATHC 70  
Db 297 CGLRQYQPOQFRKIGGLFADIASHPQAAIFVXRRSPGFRFLCGGILISSCWILSAHC 356  
Qy 71 FIDYPKKEDYIVYGRSLNSNTQGENKFEVENILHKDYSADTLAHNDIALKIRSK 130  
Db 357 FVERFPFHLLTVILGRYRVVPGEEQKFEVEKYVHKFEFDDT--YNDIALQLKSDS 414  
Qy 131 GRCAQPSRTIQTICLPSMYNDPQF---GTSCEITGFGKENSIDYLYPEQLKMTVVKLIS 186  
Db 415 SRCAQESSVVGTAQLP---DPDVLQPDWTECELSGYGKHEALSPFFSRLEKHAHVRLY 470  
Qy 197 HRCQOQPHYGEVTTMCLAADP-----QWKTDSCQSGGGLVCSLQGRMTLTGIVSW 241  
Db 471 SSRCTSQFLFKNTITSNMLCAGDTRTTCGNQDVHDACQSGGGLVCSLQGRMTLTGIVSW 530  
Qy 242 GRGQALXKPGVTVTRVSHFLPWRSHTK 270  
Db 531 GLGQKQKQVPGIYTVKTVNYLDWDRDNR 559



RESULT 10  
A29941  
C:Species: Mus musculus (house mouse)  
C:Accession: A29941  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A29941; S48205; S48207; S48206  
R:Rickles, R.J.; Darrow, A.L.; Strickland, S.  
J. Biol. Chem. 263, 1563-1569, 1988  
A:Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mRNA  
A:Reference number: A29941; MUID:88087303; PMID:2826484  
A:Accession: A29941  
A:Molecule type: mRNA  
A:Residues: 1-559 <RIC>  
A:Cross-references: GB:J03520; NID:9202109; PIDN:AAA40470.1; PID:9202110  
R:LiJinen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.  
Eur. J. Biochem. 224, 863-871, 1994  
A:Title: Characterization of the murine plasma fibrinolytic system.  
A:Reference number: S48202; MUID:95010076; PMID:7523120  
A:Accession: S48205  
A:Molecule type: protein  
A:Residues: 33-37,'X',39-40 <LIJ>  
A:Accession: S48207  
A:Molecule type: protein  
A:Residues: 309-316 <LIJ>  
A:Accession: S48206  
A:Molecule type: protein  
A:Residues: 33-37,'X',39-40 <LIJ>  
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat homology  
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F:1-17/Domain: signal sequence #status predicted <SIG>  
F:18-29/Domain: propeptide #status predicted <PRO>  
F:30-559/Product: t-plasminogen activator #status predicted <WAT>  
F:30-308/Product: t-plasminogen activator chain A #status predicted <ACH>  
F:38-75/Domain: fibronectin type I repeat homology <IFI>  
F:83-116/Domain: EGF homology <EGF>  
F:124-205/Domain: kringle homology <KR1>  
F:213-294/Domain: kringle homology <KR2>  
F:309-559/Product: t-plasminogen activator chain B #status predicted <BCH>  
F:309-553/Domain: trypsin homology <TRY>  
F:38-68,68-78,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-289,297-414/Domain: Binding site: carboxylate (Asn) (covalent) #status predicted  
F:149,481/Binding site: Arg-ile (plasma) #status predicted  
F:308-309/Cleavage site: Arg-ile (plasma) #status predicted  
F:355,404,510/Active site: His, Asp, Ser #status predicted

Query Match 37.8%; Score 569.5; DB 1; Length 559;  
Best Local Similarity 42.8%; Pred. No. 1.3e-41;  
Matches 115; Conservative 47; Mismatches 90; Indels 17; Gaps 6;  
Qy 13 CG-OKTLRPRFKIIGGFTTIENQWFAAIY-RRHRGGSVTVVCGSLISPCWVISATHC 70  
Db 297 CGLQYKRPQFRIGKGLYDITSHFWQAALFVKNKESPGERFLCGGLVLISSCWVLSAHC 356  
Qy 71 FIDYPKEDYIVLGRSLNSNTQGMKEVENLILHKDYSADTLAHHNDLALKIRSK 130  
Db 357 FLERFPNHLKVLGRVTVVPGESQTFEIKYIHBEFDDDT--IDNDIALQLRSOS 414  
Qy 131 GRCAQPSRTIQTICLPSMYNDPQF---GTSCEITGFGKNSDYLPEQLKMTVVKLIS 186  
Db 415 KQCAQESSVGTACLP---DPNLQPDWTECELSGVGKHEASSPFFSRLKNAHVRVLP 470  
Qy 187 HRECCQPHYGVSEVTTMCLAAP-----QWKTDCGDSGGGLVCSLQGRMTLITGVSW 241  
Db 471 SSRCTSHLPNKTNTNNMLCAGDTRSGNQDPLDADCGDSGGGLVCMINKQMTLTGISW 530  
Qy 242 GRGKALDKPGVTVRVSHFLPWIRSHTK 270  
Db 531 GLGGQKDVGVTVKTVNYLGDWHDNNKQ 559  
RESULT 11  
JS0599  
t-plasminogen activator (EC 3.4.21.68) beta precursor - common vampire bat  
N/Alternate names: tissue plasminogen activator

C:Species: Desmodus rotundus (common vampire bat)  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999  
C:Accession: JS0599  
R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon  
Gene 105, 229-237, 1991  
A:Title: The plasminogen activator family from the salivary gland of the vampire bat I  
A:Reference number: JS0597; MUID:92039036; PMID:1937019  
A:Accession: JS0599  
A:Molecule type: mRNA  
A:Residues: 1-431 <KRA>  
A:Cross-references: GB:M63989; NID:9166076; PIDN:AAA31594.1; PID:9166077  
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat homology  
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-36/Domain: propeptide #status predicted <PRO>  
F:37-431/Product: plasminogen activator beta #status predicted <PLA>  
F:41-74/Domain: EGF homology <EGF>  
F:82-163/Domain: kringle homology <KRG>  
F:180-425/Domain: trypsin homology <TRY>  
F:41-52,46-63,65-74,82-163,103-145,134-158,168-299,211-227,219-288,313-388/Domain: Binding site: carboxylate (Asn) (covalent) #status predicted  
F:139,352/Binding site: His-Ser (plasma) #status predicted  
F:179-180/Cleavage site: His, Asp, Ser #status predicted  
F:226,275,382/Active site: His, Asp, Ser #status predicted  
F:345-361,378-406/Disulfide bonds: #status predicted

Query Match 36.1%; Score 547; DB 2; Length 431;  
Best Local Similarity 42.3%; Pred. No. 8.4e-40;  
Matches 118; Conservative 43; Mismatches 102; Indels 12; Gaps 6  
Qy 4 SPPEELKFOCG-OKTLRPRFKIIGGFTTIENQWFAAIYRRHRGGS-VTVVCGSLISIP 61  
Db 159 SVPVCSKATCGLAKYKPEQLHSTGLFTDITSHFWQAALFVKNKESPGERFLCGGLISS 218  
Qy 62 CWVISATHCID-YPKEDYIVLGRSLNSNTQGMKEVENLILHKDYSADTLAHHND 120  
Db 219 CWLTAACHQOERYPPHQLRV-LGRTYRVKPKESQTFEIKYIHBEFDDDT--YNN 275  
Qy 121 IALIKIRSKGRCAQPSRTIQTICLPSMYNDPQFSGTSCITGFGKNSDYLPEQLKMT 180  
Db 276 IALLQLKSSGPQCAQESDVSRAICLPEANLQPLDWTCELSGVGKHEASSPFFSEQLKEG 335  
Qy 181 VKVLISRECCQPHYGVSEVTTMCLAAPQWKT-----DSQGDGGLVCSLQGRMT 234  
Db 336 HVRLYPSRSCTSKFLFNKTNTNNMLCAGDTRSGIYFNVDHACQDGGGLVCMNDNMT 395  
Qy 235 LTGIVSWGRGKALDKPGVTVRVSHFLPWIRSHTK 269  
Db 396 LLGIISWVGCGEKDIPGVTVKTVNYLGDWHDNNR 430  
RESULT 12  
JS0600  
t-plasminogen activator (EC 3.4.21.68) gamma precursor - common vampire bat  
N/Alternate names: tissue plasminogen activator  
C:Species: Desmodus rotundus (common vampire bat)  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999  
C:Accession: JS0600  
R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon  
Gene 105, 229-237, 1991  
A:Title: The plasminogen activator family from the salivary gland of the vampire bat I  
A:Reference number: JS0597; MUID:92039036; PMID:1937019  
A:Accession: JS0600  
A:Molecule type: mRNA  
A:Residues: 1-394 <KRA>  
A:Cross-references: GB:M63990; NID:9166078; PIDN:AAA31595.1; PID:9166079  
A:Note: the authors translated the codon ATC for residue 75 as Thr  
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat homology  
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-36/Domain: propeptide #status predicted <PRO>  
F:37-394/Product: plasminogen activator gamma #status predicted <PLA>  
F:45-126/Domain: kringle homology <KRG>  
F:143-388/Domain: trypsin homology <TRY>

F:45-126,66-108,97-121,131-262,174-190,182-251,276-351,308-324,341-369/Disulfide bonds:  
F:142-143/Cleavage site: His-Ser (plasmin) #status predicted  
F:189,238,345/Active site: His, Asp, Ser #status predicted  
F:315/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 36.2%; Score 546; DB 2; Length 394;  
Best Local Similarity 42.9%; Pred. No. 9.2e-39;  
Matches 118; Conservative 43; Mismatches 102; Indels 12; Gaps 6;  
  
QY 4 SPPEELKFCQG-QKTLRPRFKIIGGFTTIENQWFAAIYRHRGGS-VTVVCGSLISP 61  
DB 122 SVPVCSKATCGLRKYPQLHSTGGLFTDITSHPWQAIFAQNRSSGERFLCGILLISS 181  
  
QY 62 CWVISATHCFID-YPKKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 120  
DB 182 CWVLTAHCFQERYPPQHLRVV-LGRTYRVKPGKEQTFVEKCIIVHEEFDDT--YNN 238  
  
QY 121 IALLKIRSGRCACQPSRTIOTICLPSMYNDPQFGTSCEITGFGKNSDYLYPEQLKMT 180  
DB 239 IALLQKSGSPQCAQESDSVRAICLPEANLQPLDWTCELSGVGKHSSSPFYSEQLKEG 298  
  
QY 181 VVKLISHRECQPHYGVSVTTKMLCAADPQWKT-----DSCQDSSGGLVCSLQGRMT 234  
DB 299 HVLYPSSRCTSKFLFNKVTNNMLCAGDTRSGEIVPNVHDACQDSSGGLVCMNDNMT 358  
  
QY 235 LTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTK 269  
DB 359 LLGIISWGVGCGEKDIPGVYTKVNYLWIRDNMR 393

RESULT 13  
JS0598  
t-plasminogen activator (EC 3.4.21.68) alpha-2 precursor - common vampire bat  
N;Alternate names: tissue plasminogen activator  
C;Species: Desmodus rotundus (common vampire bat)  
C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999  
C;Accession: JS0598  
R;Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Don  
Gene 105, 229-237, 1991  
A;Title: The plasminogen activator family from the salivary gland of the vampire bat Des  
A;Reference number: JS0597; MUID:92039036; PMID:1937019  
A;Accession: JS0598  
A;Molecule type: mRNA  
A;Residues: 1-477 <KRA>  
A;Cross-references: GB:M63988; NID:gl66074; PID:AAA31593.1; PID:gl66075  
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom  
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-36/Domain: propeptide #status predicted <PRO>  
F:37-477/Product: plasminogen activator alpha-2 #status predicted <PLA>  
F:42-79/Domain: fibronectin type I repeat homology <1FA>  
F:87-120/Domain: EGF homology <EGF>  
F:128-209/Domain: kringle homology <KRG>  
F:226-471/Domain: trypsin homology <TRY>  
F:42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-345,359-4  
F:185,398/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:225-226/Cleavage site: His-Ser (plasmin) #status predicted  
F:272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 36.2%; Score 546; DB 2; Length 477;  
Best Local Similarity 42.9%; Pred. No. 1.2e-39;  
Matches 118; Conservative 43; Mismatches 102; Indels 12; Gaps 6;  
  
QY 4 SPPEELKFCQG-QKTLRPRFKIIGGFTTIENQWFAAIYRHRGGS-VTVVCGSLISP 61  
DB 205 SVPVCSKATCGLRKYPQLHSTGGLFTDITSHPWQAIFAQNRSSGERFLCGILLISS 264  
  
QY 62 CWVISATHCFID-YPKKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 120  
DB 265 CWVLTAHCFQERYPPQHLRVV-LGRTYRVKPGKEQTFVEKCIIVHEEFDDT--YNN 321  
  
QY 121 IALLKIRSGRCACQPSRTIOTICLPSMYNDPQFGTSCEITGFGKNSDYLYPEQLKMT 180  
DB 322 IALLQKSGSPQCAQESDSVRAICLPEANLQPLDWTCELSGVGKHSSSPFYSEQLKEG 381  
  
QY 181 VVKLISHRECQPHYGVSVTTKMLCAADPQWKT-----DSCQDSSGGLVCSLQGRMT 234  
DB 382 HVLYPSSRCTSKFLFNKVTNNMLCAGDTRSGEIVPNVHDACQDSSGGLVCMNDNMT 441

Query Match 36.2%; Score 546; DB 2; Length 477;  
Best Local Similarity 42.9%; Pred. No. 1.2e-39;  
Matches 118; Conservative 43; Mismatches 102; Indels 12; Gaps 6;  
  
QY 4 SPPEELKFCQG-QKTLRPRFKIIGGFTTIENQWFAAIYRHRGGS-VTVVCGSLISP 61  
DB 205 SVPVCSKATCGLRKYPQLHSTGGLFTDITSHPWQAIFAQNRSSGERFLCGILLISS 264  
  
QY 62 CWVISATHCFID-YPKKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 120  
DB 265 CWVLTAHCFQERYPPQHLRVV-LGRTYRVKPGKEQTFVEKCIIVHEEFDDT--YNN 321  
  
QY 121 IALLKIRSGRCACQPSRTIOTICLPSMYNDPQFGTSCEITGFGKNSDYLYPEQLKMT 180  
DB 322 IALLQKSGSPQCAQESDSVRAICLPEANLQPLDWTCELSGVGKHSSSPFYSEQLKEG 381  
  
QY 181 VVKLISHRECQPHYGVSVTTKMLCAADPQWKT-----DSCQDSSGGLVCSLQGRMT 234  
DB 382 HVLYPSSRCTSKFLFNKVTNNMLCAGDTRSGEIVPNVHDACQDSSGGLVCMNDNMT 441

DB 322 IALLQKSGSPQCAQESDSVRAICLPEANLQPLDWTCELSGVGKHSSSPFYSEQLKEG 381  
QY 181 VVKLISHRECQPHYGVSVTTKMLCAADPQWKT-----DSCQDSSGGLVCSLQGRMT 234  
DB 382 HVLYPSSRCTSKFLFNKVTNNMLCAGDTRSGEIVPNVHDACQDSSGGLVCMNDNMT 441  
QY 235 LTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTK 269  
DB 442 LLGIISWGVGCGEKDIPGVYTKVNYLWIRDNMR 476

RESULT 14  
A34369  
t-plasminogen activator (EC 3.4.21.68) precursor - false vampire bat (Megaderma  
C;Species: Megaderma lyra  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C;Accession: A34369  
R;Gardell, S.; Duong, L.T.; Diehl, R.E.; York, J.D.; Hare, T.R.; Register, R.; Jac  
J. Biol. Chem. 264, 17947-17952, 1989  
A;Title: Isolation, characterization, and cDNA cloning of a vampire bat saliva  
A;Reference number: A34369; MUID:90036867; PMID:2509450  
A;Accession: A34369  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-477 <GAR>  
A;Cross-references: GB:J05082; NID:gl66080; PID:AAA31596.1; PID:gl66081  
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I  
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-36/Domain: propeptide #status predicted <PRO>  
F:37-477/Product: plasminogen activator #status predicted <PLA>  
F:42-79/Domain: fibronectin type I repeat homology <1FA>  
F:87-120/Domain: EGF homology <EGF>  
F:128-209/Domain: kringle homology <KRG>  
F:226-471/Domain: trypsin homology <TRY>  
F:42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-345,359-4  
F:272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 36.1%; Score 545; DB 1; Length 477;  
Best Local Similarity 42.9%; Pred. No. 1.4e-39;  
Matches 118; Conservative 43; Mismatches 102; Indels 12; Gaps 6;  
  
QY 4 SPPEELKFCQG-QKTLRPRFKIIGGFTTIENQWFAAIYRHRGGS-VTVVCGSLISP 61  
DB 205 SVPVCSKATCGLRKYPQLHSTGGLFTDITSHPWQAIFAQNRSSGERFLCGILLISS 264  
  
QY 62 CWVISATHCFID-YPKKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 120  
DB 265 CWVLTAHCFQERYPPQHLRVV-LGRTYRVKPGKEQTFVEKCIIVHEEFDDT--YNN 321  
  
QY 121 IALLKIRSGRCACQPSRTIOTICLPSMYNDPQFGTSCEITGFGKNSDYLYPEQLKMT 180  
DB 322 IALLQKSGSPQCAQESDSVRAICLPEANLQPLDWTCELSGVGKHSSSPFYSEQLKEG 381  
  
QY 181 VVKLISHRECQPHYGVSVTTKMLCAADPQWKT-----DSCQDSSGGLVCSLQGRMT 234  
DB 382 HVLYPSSRCTSKFLFNKVTNNMLCAGDTRSGEIVPNVHDACQDSSGGLVCMNDNMT 441  
  
QY 235 LTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTK 269  
DB 442 LLGIISWGVGCGEKDIPGVYTKVNYLWIRDNMR 476

RESULT 15  
JS0597  
t-plasminogen activator (EC 3.4.21.68) alpha-1 precursor - common vampire bat  
N;Alternate names: tissue plasminogen activator  
C;Species: Desmodus rotundus (common vampire bat)  
C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jul-1999  
C;Accession: JS0597  
R;Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alago  
Gene 105, 229-237, 1991  
A;Title: The plasminogen activator family from the salivary gland of the vampi

A;Residues: 14-332,'S',334-615 <C02>  
A;Cross-references: GB:M11723; NID:g180358; PIDN:AAA51986.1; PID:g180359  
R;Que, B.G.; Davie, E.W.  
Biochemistry 25, 1525-1528, 1986  
A;Title: Characterization of a cDNA coding for human factor XII (Hageman factor).  
A;Reference number: A25191; MUID:86216049; PMID:3011063  
A;Accession: A25191  
A;Molecule type: mRNA  
A;Residues: 146-378,'G',380-615 <QUE>  
A;Cross-references: GB:M13147; NID:g180360; PIDN:AAA70224.1; PID:g180361  
R;McMullen, B.A.; Fujikawa, K.  
J. Biol. Chem. 260, 5328-5341, 1985  
A;Title: Amino acid sequence of the heavy chain of human alpha-factor XIIa (activated factor XIIa).  
A;Reference number: A22248; MUID:85182674; PMID:3886654  
A;Accession: A22248  
A;Molecule type: Protein  
A;Residues: 20-379 <MCW>  
R;Fujikawa, K.; McMullen, B.A.  
J. Biol. Chem. 258, 10924-10933, 1983  
A;Title: Amino acid sequence of human beta-factor XIIa.  
A;Reference number: A21037; MUID:83291041; PMID:6604055  
A;Accession: A21037  
A;Molecule type: Protein  
A;Residues: 354-362;373-615 <FUD>  
R;Harris, R.J.; Ling, V.T.; Spellman, M.W.  
J. Biol. Chem. 267, 5102-5107, 1992  
A;Title: O-linked fucose is present in the first epidermal growth factor domain of factor XIIa.  
A;Reference number: A44605; MUID:92184750; PMID:1544894  
A;Contents: annotation; carbohydrate binding site  
C;Genetics:  
A;Gene: GDB:F12  
A;Cross-references: GDB:119892; OMIM:234000  
A;Map position: 5q34-5qter  
A;Introns: 19/3; 39/1; 72/2; 96/1; 133/1; 177/1; 212/1; 267/2; 340/1; 417/2; 451/1; 511/1  
C;Complex: factor XII, prekallikrein, and HWW kininogen form a complex bound to plasminogen.  
C;Function:  
A;Description: factor XIIa catalyzes the proteolytic activation of plasminogen.  
ikrein  
A;Pathway: blood coagulation; fibrinolysis  
C;Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat  
C;Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; kringle;  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-372,373-615/Product: coagulation factor XIIa, alpha form #status experiment  
F;47-88/Domain: fibronectin type II repeat homology <FB2>  
F;98-130/Domain: EGF homology <EGL>  
F;135-170/Domain: fibronectin type I repeat homology <LF1>  
F;178-209/Domain: EGF homology <EG2>  
F;217-295/Domain: kringle homology <KR>  
F;298-356/Region: proline-rich  
F;354-362,373-615/Product: coagulation factor XIIa, beta form #status experiment  
F;373-609/Domain: trypsin homology <TRY>  
F;98-110,104-119,121-130,135-163,161-170,178-189,183-198,200-209,217-295,238-266-267-268-269-270-271-272-273-274-275-276-277-278-279-280-281-282-283-284-285-286-287-288-289-290-291-292-293-294-295-296-297-298-299-300-301-302-303-304-305-306-307-308-309-310-311-312-313-314-315-316-317-318-319-320-321-322-323-324-325-326-327-328-329-330-331-332-333-334-335-336-337-338-339-340-341-342-343-344-345-346-347-348-349-350-351-352-353-354-355-356-357-358-359-360-361-362-363-364-365-366-367-368-369-370-371-372-373-374-375-376-377-378-379-380-381-382-383-384-385-386-387-388-389-390-391-392-393-394-395-396-397-398-399-400-401-402-403-404-405-406-407-408-409-410-411-412-413-414-415-416-417-418-419-420-421-422-423-424-425-426-427-428-429-430-431-432-433-434-435-436-437-438-439-440-441-442-443-444-445-446-447-448-449-450-451-452-453-454-455-456-457-458-459-460-461-462-463-464-465-466-467-468-469-470-471-472-473-474-475-476-477-478-479-480-481-482-483-484-485-486-487-488-489-490-491-492-493-494-495-496-497-498-499-500-501-502-503-504-505-506-507-508-509-510-511-512-513-514-515-516-517-518-519-520-521-522-523-524-525-526-527-528-529-530-531-532-533-534-535-536-537-538-539-540-541-542-543-544-545-546-547-548-549-550-551-552-553-554-555-556-557-558-559-560-561-562-563-564-565-566-567-568-569-570-571-572-573-574-575-576-577-578-579-580-581-582-583-584-585-586-587-588-589-590-591-592-593-594-595-596-597-598-599-600-601-602-603-604-605-606-607-608-609-610-611-612-613-614-615-616-617-618-619-620-621-622-623-624-625-626-627-628-629-630-631-632-633-634-635-636-637-638-639-640-641-642-643-644-645-646-647-648-649-650-651-652-653-654-655-656-657-658-659-660-661-662-663-664-665-666-667-668-669-670-671-672-673-674-675-676-677-678-679-680-681-682-683-684-685-686-687-688-689-690-691-692-693-694-695-696-697-698-699-700-701-702-703-704-705-706-707-708-709-710-711-712-713-714-715-716-717-718-719-720-721-722-723-724-725-726-727-728-729-730-731-732-733-734-735-736-737-738-739-740-741-742-743-744-745-746-747-748-749-750-751-752-753-754-755-756-757-758-759-760-761-762-763-764-765-766-767-768-769-770-771-772-773-774-775-776-777-778-779-780-781-782-783-784-785-786-787-788-789-790-791-792-793-794-795-796-797-798-799-800-801-802-803-804-805-806-807-808-809-810-811-812-813-814-815-816-817-818-819-820-821-822-823-824-825-826-827-828-829-830-831-832-833-834-835-836-837-838-839-840-841-842-843-844-845-846-847-848-849-850-851-852-853-854-855-856-857-858-859-860-861-862-863-864-865-866-867-868-869-870-871-872-873-874-875-876-877-878-879-880-881-882-883-884-885-886-887-888-889-890-891-892-893-894-895-896-897-898-899-900-901-902-903-904-905-906-907-908-909-910-911-912-913-914-915-916-917-918-919-920-921-922-923-924-925-926-927-928-929-930-931-932-933-934-935-936-937-938-939-940-941-942-943-944-945-946-947-948-949-950-951-952-953-954-955-956-957-958-959-960

QY 174 PEOLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSGCGDGGPLVCSLOG-- 231  
 Db 516 ASFLQBAQVFFSLERCSAPDVHGSSILPGLM/CAGFEGGTACQCGDGGPLVCSQDAE 575  
 QY 232 -RMTLTGIVSWGRGCALKDKPGVYTVSVHFLPWIRSH 268  
 Db 576 RRLTQGLIISWGGCGDRNKPQVYTDVAYLAWIREHT 613

## RESULT 17

A46688  
 hepatocyte growth factor activator (EC 3.4.21.-) precursor [validated] - human  
 C:Species: Homo sapiens (man)  
 C>Date: 21-Sep-1993 #sequence\_revision 25-Aug-1995 #text\_change 08-Dec-2000  
 C:Accession: A46688  
 R:Wiyazawa, K.; Shimomura, T.; Kitamura, A.; Kondo, J.; Morimoto, Y.; Kitamura, N.  
 J. Biol. Chem. 288, 10024-10028, 1993  
 A>Title: Molecular cloning and sequence analysis of the cDNA for a human serine protease  
 d coagulation factor XII.  
 A:Reference number: A46688; MUID:93252878; PMID:7683665  
 A:Accession: A46688  
 A:Molecule type: mRNA  
 A:Residues: 1-655 <M1>  
 A:Cross-references: DDBJ:D14012; MID:g219680; PIDN:BAA03113.1; PID:g219681.  
 A:Experimental source: liver (mRNA); serum (protein)  
 A:Note: sequence extracted from NCBI backbone (NCBI:131227, NCBIP:131228)  
 A:Note: parts of the sequence, including the amino ends of the heavy and light chains, c  
 C:Genetics:  
 A:Gene: GDB:HGPAC; HGFA: HGPAP  
 A:Cross-references: GDB:9954514  
 A:Map position: 4p16-4p16  
 C:Function:  
 A:Description: activates hepatocyte growth factor by specific proteolytic cleavage  
 A:Pathway: tissue repair and regeneration  
 C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;  
 C:Keywords: glycoprotein; hydrolase; kringle; liver; plasma; serine proteinase  
 F:1-34/Domain: signal sequence #status predicted <SIG>  
 F:108-148/Domain: fibronectin type II repeat homology <IF2>  
 F:164-197/Domain: EGF homology <EG1>  
 F:202-237/Domain: fibronectin type I repeat homology <IF1>  
 F:245-278/Domain: EGF homology <EG2>  
 F:286-367/Domain: kringle homology <KRG>  
 F:373-407/Product: hepatocyte growth factor activator light chain #status experimental <  
 F:408-655/Product: hepatocyte growth factor activator heavy chain #status experimental <  
 F:408-641/Domain: trypsin homology <TRY>  
 F:408-48,290,468,492,546/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:164-175,169-186,188-197,202-230,228-237,245-256,250-267,269-278,286-367,307-349,338-36  
 F:447,497,598/Active site: His, Asp, Ser #status predicted

Query Match 32.7%; Score 493.5; DB 1; Length 655;

Best Local Similarity 38.3%; Pred. No. 6.2e-35;

Matches 105; Conservative 43; Mismatches 109; Indels 17; Gaps 6;

QY 1 KPSPPEELKPCQCK-----TLRPKLIIGGFTTIENOPFAALYRRHRGSGVTVCG 55  
 Db 385 EPASPGRQ---ACGRHKKRTFLPR---IIGSSSLSPGHPMLAAIY---IGDS---FCA 433  
 QY 56 GSLISPCWISATCFIDYPKEDYIVYLGSRINSNTQGMKFEVENLILHKDYADTL 115  
 Db 434 GSVHTCWVSAACFSESPRDSVSVLQGFNFRTDVTQFGIEKIPYLYSVFNP 493  
 QY 116 AHNDIALKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCITGFGKENSTDYLYPE 175  
 Db 494 SDH-DLVLRLLKGRDCATRSQVQPCILPEPGSTFPAGHKCQIAGWGLHLDNVSGYSS 552  
 QY 176 QLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSGCGDGGPLVCSLQGM 235  
 Db 553 SLREALVPLVADHKCSSPEVVGADISPNMLCAGVDFCKSDACQDGGPLACENGVAYL 612  
 QY 236 TGIYSWGRGCALKDKPGVYTVSVHFLPWIRSH 269  
 Db 613 YGIIISWGGCGGLRHKPGVYTVSVHFLPWIRSH 646

## RESULT 18

S28941  
 coagulation factor XIIa (EC 3.4.21.38) - guinea pig (fragment)  
 N:Alternate names: Hageman factor  
 C:Species: Cavia porcellus (guinea pig)  
 C>Date: 25-Feb-1994 #sequence\_revision 03-Aug-1995 #text\_change 21-Jan-2000  
 C:Accession: S28941  
 R:Semba, U.; Yamamoto, T.; Kunisada, T.; Shibuya, Y.; Tanase, S.; Kambara, T.;  
 Biochim. Biophys. Acta 1159, 113-121, 1992  
 A>Title: Primary structure of guinea-pig Hageman factor: sequence around the c  
 A:Reference number: S28941; MUID:93003367; PMID:1390917  
 A:Accession: S28941  
 A:Molecule type: mRNA  
 A:Status: preliminary  
 A:Residues: 1-603 <SM>  
 A:Cross-references: EMBL:X68615; NID:g49578; PIDN:CAA48600.1; PID:g49579  
 C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat  
 C:Keywords: hydrolase; serine proteinase  
 F:46-87/Domain: fibronectin type II repeat homology <IF2>  
 F:134-169/Domain: fibronectin type I repeat homology <IF1>  
 F:177-208/Domain: EGF homology <EGF>  
 F:216-294/Domain: kringle homology <KRG>  
 F:359-597/Domain: trypsin homology <TRY>

Query Match 32.2%; Score 485.5; DB 2; Length 603;

Best Local Similarity 40.7%; Pred. No. 2.8e-34;

Matches 111; Conservative 41; Mismatches 102; Indels 19; Gaps 8

QY 6 PEELKFCQCKTLRPRF---KIIGGFTTIENOPFAALYRRHRGSGVTVCGSLISP 61  
 Db 338 PETSLLCCQR-LRKRLSSLSRIVGLVALPGAHPYIAALY---WGS--NFCSGSLIAP 390  
 QY 62 CWVISATHCFIDYPKEDYIVYLGSRINSNTQGMKFEVENLILHKDYADTLAHNDI 121  
 Db 391 CWLTAARCLQRRPAPBELKVLQGRHNSCEHCQTLAVHSYRLHEAFSPS--SYLNDL 448  
 QY 122 ALLKT-RSEGRCAQPSRTIQTICLPSMYNDPQFG--TSCBITGFGKENSTDYLYPQLK 178  
 Db 449 ALLRLQKSDGSCAQLPYQVTVCLPSGAPPSESETTCCEVAGWGHQFSGAEYSFQLQ 508  
 QY 179 MTVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSGCGDGGPLVC---SLQGRMTL 235  
 Db 509 EAQVPLISSERCSSPEVVGADFLSGMLCAGFLEGTDACQDGGPLVCEDEAAEHLIL 568  
 QY 236 TGIYSWGRGCALKDKPGVYTVSVHFLPWIRSH 268  
 Db 569 RGIYSWGGCGGRNRPQVITDVASILTQIKHT 601

## RESULT 19

S45281

coagulation factor XIIa (EC 3.4.21.38) precursor - bovine (fragment)

N:Alternate names: Hageman factor (activated)

C:Species: Bos primigenius taurus (cattle)

C&gt;Date: 10-Apr-1995 #sequence\_revision 22-Apr-1995 #text\_change 21-Jan-2000

C:Accession: S45281; A61329

R:Shibuya, Y.; Semba, U.; Okabe, H.; Kambara, T.; Yamamoto, T.

Biochim. Biophys. Acta 1206, 63-70, 1994

A&gt;Title: Primary structure of bovine Hageman factor (blood coagulation factor

A:Reference number: S45281; MUID:94242782; PMID:8186251

A:Accession: S45281

A:Molecule type: mRNA

A:Residues: 1-593 &lt;SH1&gt;

A:Cross-references: GB:S70164

A:Note: the authors translated the codon GAG for residue 23 as Val, GAG for re-

s Pro. CTC for residue 203 as Phe, GTG for residue 247 as Leu, CCG for residue

is, and ATC for residue 505 as Leu

R:Fujikawa, K.; Walsh, K.A.; Davie, E.W.

Biochemistry 16, 2270-2278, 1977

A&gt;Title: Isolation and characterization of bovine factor XII (Hageman factor).

A:Reference number: A61329; MUID:77182112; PMID:861210

555 CWLTTAAHCFKRYGNNRSYAVRVDYHTLVPEEFEEQETCVQOIVHNRYPDRSDY--D 612

J. Biochem. 119, 1157-1165, 1996

555 CWJLTAACHCEKRYGNNSSRSVAVRUGDYHTI VPEEEFEOETGVCOOTVTHRNYPBDSY--D 613

A;Title: Purification and characterization of a novel hyaluronan-binding protein (PHBP)  
r activator.  
A;Reference number: JC4795; MUID:96425001; PMID:8827452  
A;Accession: JC4795  
A;Molecule type: mRNA  
A;Residues: 1-560 <CHO>  
A;Cross-references: GB:883182; NID:G1836158; PIDN:ABA46909.1; PID:G1836159  
A;Experimental source: plasma  
A;Note: parts of this sequence, including the amino ends of the mature chains, were determined by protein sequencing  
C;Genetics:  
A;Gene: GDB:HABP2; HABP; PHBP; HGPAL  
A;Cross-references: GDB:1457392  
C;Complex: a disulfide-bonded heterodimer of chains produced from the same precursor; the heavy chain is a disulfide-bonded heterodimer of chains produced from the same precursor; the light chain is a disulfide-bonded heterodimer of chains produced from the same precursor  
C;Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology; trypsin homology; chondroitin sulfate proteoglycan; glycoprotein; hyaluronate; hyaluronidase; hyaluronidase  
F;1-23/Domain: signal sequence #status predicted <SIG>  
F;24-313/Product: plasma hyaluronan-binding protein, 50K chain #status predicted <50K>  
F;77-108/Domain: EGF homology <EG1>  
F;115-147/Domain: EGF homology <EG2>  
F;154-187/Domain: EGF homology <EG3>  
F;194-276/Domain: kringle homology <KR1>  
F;314-550/Domain: trypsin homology <TRY>  
F;314-516/Product: plasma hyaluronan-binding protein, catalytic chain #status predicted  
F;54-207/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;77-88, 82-97, 99-108, 115-125, 120-136, 138-147, 154-165, 159-176, 178-187, 194-276, 215-257, 246-362, 405-509/Active site: His, Asp, Ser #status predicted  
Query Match 27.7%; Score 417; DB 1; Length 560;  
Best Local Similarity 36.2%; Pred. No. 2.3e-28;  
Matches 104; Conservative 40; Mismatches 113; Indels 30; Gaps 11;  
QY 2 PSSPPBELK--FQCGKTLRPR--FKLIGGEFTTNPQWFAAIYRHRGGSVT-----51  
Db 288 PTPSTKLPGDSCGTEAERKIKRIYGGFKSTAGKHPQASLQ-----SSLPLTISMP 342  
QY 52 --YVCGSLSPCWISATHCFIDYPKEDYIVLGRSLNNTGEMKFEVENLILHKD 109  
Db 343 QGHFCGALLHPCWLTAHC-TDI-KTRHLKVLGDQDLKKEEHEQSFVEKFKYSH 400  
QY 110 YSADTLAHNDIALKIRSKGCAQPSRTIQTICLPSMYNDPQF--GTSCBITGFKEN 167  
Db 401 YNERDEIPHNDIALKLPVDGRCALSKYKVTCLP-----DGSFPGSECHISGWVTE 456  
QY 168 STDLYPEQLKMTVWKLISHRECCQPHYVGSVETTKMLCAADPQWK-TDSCGDSGGPLV 226  
Db 457 TKG--GSRQLLDKVLKLIANTLNSQLVDHMDSDNLCAGNLQKPGQDTGDSGGPLT 514  
QY 227 CSLOGRMTLGIWSRGCAKDKPGYTRVSHFLPWSHTEKNG 273  
Db 515 CERDGTYYVYGVISWGLECG--KRPVYTVTKFLNWKATIKSESG 559  
RESULT 23  
A57014  
proctasin (EC 3.4.21.-) precursor - human  
C;Species: Homo sapiens (man)  
C;Date: 24-May-1996 #sequence\_revision 24-May-1996 #text\_change 21-Apr-2003  
C;Accession: A57014; A54866  
R;Yu, J. X.; Chao, L.; Chao, J.  
J. Biol. Chem. 270, 13483-13489, 1995  
A;Title: Molecular cloning, tissue-specific expression, and cellular localization of human proctasin  
A;Reference number: A57014; MUID:95286644; PMID:7768952  
A;Accession: A57014  
A;Status: translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-343 <RES>  
A;Cross-references: GB:141351; NID:G862304; PIDN:AA41759.1; PID:G862305  
A;Experimental source: prostate  
A;Note: parts of this sequence were determined by protein sequencing  
R;Yu, J. X.; Chao, L.; Chao, J.  
J. Biol. Chem. 269, 18843-18848, 1994  
A;Title: Proctasin is a novel human serine proteinase from seminal fluid. Purification,  
A;Reference number: A54866; MUID:94308140; PMID:8034638

A;Accession: A54866  
A;Molecule type: protein  
A;Residues: 45-64 <YUA>  
C;Genetics:  
A;Gene: GDB:PRSS8  
A;Cross-references: GDB:676446; OMIM:600823  
A;Map position: 16p11.2-16p11.2  
C;Superfamily: trypsin; trypsin homology  
C;Keywords: glycoprotein; hydrolase; serine proteinase; transmembrane protein  
F;1-32/Domain: signal sequence #status predicted <SIG>  
F;33-44, 45-343/Product: proctasin #status predicted <MAT>  
F;33-44/Domain: proctasin light chain #status predicted <CHL>  
F;45-343/Domain: proctasin heavy chain #status predicted <CHH>  
F;45-281/Domain: trypsin homology <TRY>  
F;323-341/Domain: transmembrane #status predicted <TM1>  
F;37-154, 70-86, 168-244, 201-223, 234-262/disulfide bonds: #status predicted  
F;185, 134, 238/Active site: His, Asp, Ser #status predicted  
F;155/Binding site: carbohydrate (Asn) (covalent) #status experimental  
Query Match 27.1%; Score 408.5; DB 1; Length 343;  
Best Local Similarity 36.6%; Pred. No. 7.1e-28;  
Matches 100; Conservative 36; Mismatches 100; Indels 37; Gaps 8  
QY 13 CGQKTLRPRKILGGEBTTTNPQWFAAIYRHRGGSVT-----VCGSLSPCWISAT 68  
Db 37 CG--VAPOARITGSSAVAGQWPQV-----SITVEGVHVCSSVSEQWVLSAA 84  
QY 69 HCFIDYPKEDYIVLGRSLNNTGEMKFEVENLILHKDYSADTLAHNDIALKIRS 128  
Db 85 HCFPSEHKEAYEVLGAHQDLSDYSEDAKVSTLKDILPHPSYLOE--GSGDIALLOU-- 140  
QY 129 KEGCAQPSRTIQTICLPSMYNDPQFQTSCEITFGFK-ENSTDYLYPEQLKMTVWKLISH 187  
Db 141 --SRPITFSYRIRPILCPAANASFPNGLHCTVGTGHWAPSVSLTLPKPIQQLVEPLISR 198  
QY 188 REC-----QQPHYVGSVETTKMLCAADPQWKTDSCGDSGGPLVCSLOGRMTLTG 237  
Db 199 ETNCLNINIDAKPEEPHF----VQEDNVCAGYVEGKDACQDGGSGPLSCFVSLWYLTG 254  
QY 238 IVSWGRGCAKDKPGYTRVSHFLPWSHTEK 270  
Db 255 IVSWGDAAGARNRPGVYTVASSYASMTQSKVTE 287  
RESULT 24  
KQMSPL  
plasma kallikrein (EC 3.4.21.34) precursor - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 18-Jun-1999  
C;Accession: A36557  
R;Seidah, N.G.; Sawyer, N.; Hamelin, J.; Mion, P.; Beaubien, G.; Brachpapa, L.  
DNA Cell Biol. 9, 737-748, 1990  
A;Title: Mouse plasma kallikrein: cDNA structure, enzyme characterization, and  
A;Reference number: A36557; MUID:91090844; PMID:2264928  
A;Accession: A36557  
A;Molecule type: mRNA  
A;Residues: 1-638 <SRI>  
A;Cross-references: GB:M58588; NID:G200358; PIDN:AAA63393.1; PID:G200359  
A;Note: part of this sequence, including the amino ends of both the heavy and  
C;Comment: This protein, synthesized in the liver, circulates as a noncovalent  
are linked by one or more disulfide bonds.  
C;Superfamily: coagulation factor XI; trypsin homology  
C;Keywords: blood coagulation; duplication; fibrinolysis; glycoprotein; hydrol  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-390/Product: plasma kallikrein heavy chain #status experimental <CH>  
F;20-109/Domain: apple repeat <AP1>  
F;110-199/Domain: apple repeat <AP2>  
F;200-289/Domain: apple repeat <AP3>  
F;291-380/Domain: apple repeat <AP4>  
F;391-638/Product: plasma kallikrein light chain #status experimental <LCH>  
F;391-621/Domain: trypsin homology <TRY>  
F;21-104, 47-77, 51-57, 111-194, 137-166, 141-147, 201-284, 227-256, 231-237, 292-375, 3



F:127,215,308,396,494/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:434,483,578/Active site: His, Asp, Ser #status predicted

Query Match 27.0%; Score 407; DB 1; Length 638;  
Best Local Similarity 33.8%; Pred. No. 28-27;  
Matches 89; Conservative 55; Mismatches 103; Indels 16; Gaps 6;

Qy 17 TLPRFKIIGBETTNIENQWFAAIVRRHGGSVTVVCGSLTSPCWISATHCFIDYK 76  
Db 384 TTKINARIWGGTNASLGEWQVSL--QVQLVSTHLCGSGTIGRQWVLTAAHCFDGPY 441  
Qy 77 KEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHNDIALKIRSKGRCQAP 136  
Db 442 PDWRIYGGILSLSITETTPSRKELIHQEVKYS--EGNYDIALIKLQFP---LNY 495  
Qy 137 SRTIQTICLPSMYNDPQFGTSCETGFG---KENSTDYLYPQLKMTVVKLISHRECOOP 193  
Db 496 TEFQKICLPSKADNTNITVNCWVGWTKGQGTQNT----LQKATIPLVNEECQK- 550  
Qy 194 HYVSGSVTTKMLCAADPQNKTDSCQDGGPLVCSLQGRMTLTGTVSNGRCALDKPGV 253  
Db 551 KYRDYVINKQMICAGYKEGTDACKDGGSGPLVCKHSGRWQLVGTSGEGGRKQPGV 610  
Qy 254 YTRVSHFLPWIRSHTKENGLAL 276  
Db 611 YTKVSEYMDWILEKTQSSDVAL 633

RESULT 25

KQHP

Plasma kallikrein (EC 3.4.21.34) precursor - human  
N;Alternate names: Kininogenin; plasma prekallikrein  
C;Species: Homo sapiens (man)  
C;Date: 13-Aug-1986 #sequence\_revision 13-Aug-1986 #text\_change 18-Jun-1999  
C;Accession: A00921; A37939  
R;Chung, D.W.; Fujikawa, K.; McMullen, B.A.; Davie, E.W.  
Biochemistry 25, 2410-2417, 1986  
A;Title: Human plasma prekallikrein, a zymogen to a serine protease that contains four  
A;Reference number: A00921; PMID:86243359; PMID:3521732  
A;Accession: A00921  
A;Molecule type: mRNA  
A;Residues: 1-638 <CHU>  
A;Cross-references: GB:M13143; NID:G190262; PIDN:AAA60153.1; PID:G190263  
R;McMullen, B.A.; Fujikawa, K.; Davie, E.W.  
Biochemistry 30, 2050-2056, 1991  
A;Title: Location of the disulfide bonds in human plasma prekallikrein: the presence of  
A;Reference number: A37939; PMID:91152016; PMID:1998666  
A;Accession: A37939  
A;Molecule type: protein  
A;Residues: 20-27;40-46, 'X', 48, 'H', 50, 'X', 52-70, 'H', 75-76, 'X', 78-80;103;113;131-140;141-  
260-283; 'X', 285-287, 291, 'X', 293-295;314-317, 'X', 319-320;321-324; 'X', 328-333;334-339, 'X',  
525;538-551;562, 'X', 564-567;573, 'X', 575-576;578-583, 'X', 585;592-604 <WCM>  
C;Comment: This protein, synthesized in the liver, circulates as a noncovalent complex w  
C;Comment: The zymogen is activated by factor XIIa, which cleaves the molecule into a li  
are linked by one or more disulfide bonds.  
C;Comment: The enzyme cleaves Lys-Arg and Arg-Ser bonds. It activates, in a reciprocal  
inogen and may also play a role in the renin-angiotensin system by converting prorenin i  
C;Genetics:  
A;Gene: GDB:KUK3  
A;Cross-references: GDB:I27575; OMIM:229000  
A;Map position: 4q35-4q35  
C;Superfamily: coagulation factor XI; trypsin homology  
C;Keywords: blood coagulation; duplication; fibrinolysis; glycoprotein; hydrolase; infla  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-638/Product: plasma kallikrein #status predicted <MAT>  
F:20-390/Domain: plasma kallikrein heavy chain #status predicted <HCH>  
F:20-109/Domain: apple repeat <API>  
F:110-199/Domain: apple repeat <AP2>  
F:200-289/Domain: apple repeat <AP3>  
F:291-380/Domain: apple repeat <AP4>  
F:391-638/Domain: plasma kallikrein light chain #status predicted <LCH>  
F:391-621/Domain: trypsin homology <TRY>  
F:21-104;47-77;51-57;111-194;137-166;141-147,201-284,227-256,231-237,292-375,322-328,383

F:127,308,396,453,494/Binding site: carbohydrate (Asn) (covalent) #status expe  
F:318-347;340-345/Disulfide bonds: #status predicted  
F:390-391/Cleavage site: Arg-Ile (coagulation factor XIIa) #status predicted  
F:434,483,578/Active site: His, Asp, Ser #status predicted

Query Match 26.7%; Score 402; DB 1; Length 638;  
Best Local Similarity 33.1%; Pred. No. 54e-27;  
Matches 88; Conservative 56; Mismatches 98; Indels 24; Gaps 7

Qy 17 TLPRFKIIGBETTNIENQWFAAIVRRHGGSVTVVCGSLTSPCWISATHCF 71  
Db 384 TTKTSTRIVGTTSSMGWEPQVSLQVQLTAQSH-----LCGGSLIGHQWVLTAAHCF 436  
Qy 72 IDVPKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHNDIALKIRSKEG 131  
Db 437 DGLPQDVWRIYSGILNLSITKDTFSPQIKELIIHQYKVSQNH--DIALIKLQAP-- 492  
Qy 132 RCAQPSRTIQTICLPSMYNDPQFGTSCETGFG---KENSTDYLYPQLKMTVVKLISHRE 189  
Db 493 --LNYTEFQKICLPSKADNTNITVNCWVGWTKGQGTQNT----LQKATIPLVNEE 547  
Qy 190 CQPHYVSGSVTTKMLCAADPQNKTDSCQDGGPLVCSLQGRMTLTGTVSNGRCALDK 249  
Db 548 CQK--RYQDYKITQRMVVCAGYKEGTDACKDGGSGPLVCKHSGRWQLVGTSGEGGR 606  
Qy 250 KPGVYTRVSHFLPWIRSHTKENGLA 275  
Db 607 QPGVYTKVSEYMDWILEKTQSSDGA 632

RESULT 26

S00845  
hepsin (EC 3.4.21.-) - human  
C;Species: Homo sapiens (man)  
C;Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 18-Jun-1999  
C;Accession: S00845  
R;Leytus, S.P.; Loeb, K.R.; Hagen, F.S.; Kurachi, K.; Davie, E.W.  
Biochemistry 27, 1067-1074, 1988  
A;Title: A novel trypsin-like serine protease (hepsin) with a putative transme  
A;Reference number: S00845; PMID:88209431; PMID:2835076  
A;Accession: S00845  
A;Molecule type: mRNA  
A;Residues: 1-417 <LSV>  
A;Cross-references: EMBL:X07732; NID:G32063; PIDN:CAA30558.1; PID:G32064  
C;Genetics:  
A;Gene: GDB:HPN; TMPRSS1; hepsin  
A;Cross-references: GDB:I35685; OMIM:142440  
A;Map position: 19q11-19q13.2  
C;Superfamily: hepsin; trypsin homology  
C;Keywords: hydrolase; liver; serine proteinase; transmembrane protein  
F:23-45/Domain: transmembrane #status predicted <TM>  
F:163-400/Domain: trypsin homology <TRY>  
F:188-204,291-359,322-338,349-381/Disulfide bonds: #status predicted  
F:203,257,353/Active site: His, Asp, Ser #status predicted

Query Match 26.3%; Score 396.5; DB 1; Length 417;  
Best Local Similarity 35.8%; Pred. No. 9.8e-27;  
Matches 101; Conservative 51; Mismatches 89; Indels 41; Gaps 14

Qy 13 CQKTLRPRFKIIGBETTNIENQWFAAIVRRHGGSVTVVCGSLTSPCWISATHCFI 72  
Db 153 CGRKRL-PVDRIVGRDTSIGRWFPQVSL--RYDG---AHLCCGSLSGDWVLTAAHCF- 205  
Qy 73 DYPKEDYIVYLGSRSLNSNTQGS-----MKFEVENILHKDY---SADTLAHNDIAL 123  
Db 206 --PERNR---VLSRWRFAGAVQAQSPHGLQLGVQAVVYHGGVLPFRDPHSEENSNDIAL 260  
Qy 124 LKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCETGFGKFNSTDYLYPEQ---LKMT 180  
Db 261 VHLSSP---LPLTEYIQVCLPAAQALVDGKICTVTGWG---NTQY-YGQAQGLQEA 312  
Qy 191 VVKLISHRECOQPHYVSGSVTTKMLCAADPQNKTDSCQDGGSGPLVCKHSGRWQLVGTSGEGGR 236

```
Db 313 RVPILISNDVCGADFYGNQIKPMFCAGYPEGGIDACQDGGPFCVCEDSISRTPRWRLC 372
Qy 237 GIVSWGRCALKDKPGVYTRVSHFLPW---IRSHTEENGL 274
Db 373 GIVSWGRCALKDKPGVYTRVSHFLPW---IRSHTEENGL 274
Db 373 GIVSWGRCALKDKPGVYTRVSHFLPW---IRSHTEENGL 274

RESULT 27
S55066
trypsin (EC 3.4.21.4) II precursor, pancreatic (clone 2-P29) - chicken
N:Alternate names: trypsinogen II
C:Species: Gallus gallus (chicken)
C:Date: 23-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 22-Jun-1999
C:Accession: S55066; S72347
R:Wang, K.; Gan, L.; Lee, I.; Hood, L.
Biochem. J. 307, 471-479, 1995
A:Title: Isolation and characterization of the chicken trypsinogen gene family.
A:Reference number: S55065; MUID:95251611; PMID:7733885
A:Accession: S55066
A:Molecule type: mRNA
A:Residues: 1-248 <WANI>
A:Cross-references: EMBL:U15157; NID:G603906; PID:AAA79914.1; PID:G603907
A:Experimental source: clone 2-P29
A:Accession: S72347
A:Molecule type: DNA
A:Residues: 1-248 <WANI>
A:Cross-references: EMBL:U15157; NID:G603906; PID:AAA79914.1; PID:G603907
A:Experimental source: clone 2-P29
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-25/Domain: activation peptide #status predicted <APT>
F:26-248/Product: trypsin II #status predicted <MAT>
F:26-241/Domain: trypsin homology <TRY>
F:65,109,202/Active site: His, Asp, Ser #status predicted

Query Match 26.2%; Score 395.5; DB 2; Length 248;
Best Local Similarity 38.7%; Pred. No. 6.5e-27;
Matches 94; Conservative 40; Mismatches 84; Indels 25; Gaps 7;

Qy 23 KIIGGFTTTEENQWFAAIYRRHGGSVTVVCGSLISPCWVLSATHCFIDYPKEDYIV 82
Db 25 KIVGGYTCPEHSPYQVSL-----NSGYHFCGSLINSQWVLSAHCY-----KSGIQV 73
Qy 83 YLGRSLRNSNTQGMKEFEVENLILHKDYSADTLAHNDIALLKIRSGECRCAQPSRTIQT 142
Db 74 RLGEYNDVQEDSEVVRSSVIRHPKYSITL--NNDIMLIKLS---AVEYSADIQ 127
Qy 143 ICLPSMNDPQFGTSCITGFKENSTDYLYPEOLKMTVVKLISHRECQPHYVGSVTT 202
Db 128 IALPS--SCAKAGTECLISGWNTLSNGYNYPELLOCLNAPILSDQCEA--YPGDITS 183
Qy 203 KMLCAADPOWKTDSQGDGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLP 262
Db 184 NMICVGFLEGGKDSQGDGGPVCNGE---LQGIWSWIGCALKGYPGVYTKVCNYVD 239
Qy 263 WIR 265
Db 240 WIQ 242

RESULT 28
TREOTR
trypsin (EC 3.4.21.4) precursor - bovine
N:Contains: trypsinogen
C:Species: Bos primigenius taurus (cattle)
C:Date: 24-Apr-1984 #sequence_revision 28-Feb-1986 #text_change 18-Jul-1997
C:Accession: A90164; A00946; S08774
R:Mikes, O.; Holeysovsky, V.; Tomasek, V.; Sorm, F.
Biochem. Biophys. Res. Commun. 24, 346-352, 1966
A:Title: Covalent structure of bovine trypsinogen. The position of the remaining amides.
A:Reference number: A90164; MUID:67168848; PMID:5967094
A:Accession: A90164
```

```
A:Molecule type: protein
A:Residues: 1-57,'Q',59-150,'N',152-176,'N',178-229 <MIK>
R:Hartley, E.S.
Philos. Trans. R. Soc. Lond. B257, 77-87, 1970
A:Reference number: A93755
A:Contents: annotation; revisions
R:Titani, K.; Ericsson, L.H.; Neurath, H.; Walsh, K.A.
Biochemistry 14, 1358-1366, 1975
A:Title: Amino acid sequence of dogfish trypsin.
A:Reference number: A00950; MUID:75146445; PMID:1092332
A:Contents: annotation; revisions
A:Note: the sequence agrees with that shown
R:Boyd, W.; Schwager, P.
J. Mol. Biol. 98, 693-717, 1975
A:Title: The refined crystal structure of bovine beta-trypsin at 1.8 angstrom
A:Reference number: A92954; MUID:75072097; PMID:542
C:Contents: annotation; X-ray crystallography; binding sites for calcium, subs
C:Comment: Trypsinogen is synthesized in the acinar cells of the pancreas.
C:Comment: Autocatalytic cleavage after Lys-6 leads to beta-trypsin by releasi
s pseudotrypsin. A cleavage may also occur after Arg-105.
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
F:1-229/Product: trypsinogen #status experimental <ZYM>
F:1-6/Domain: activation peptide #status experimental <APT>
F:7-222/Domain: trypsin homology <TRY>
F:7-131,132-229/Product: alpha-trypsin #status experimental <MPT>
F:6-7/Cleavage site: Lys-Ile (enteropeptidase) #status experimental
F:13-143,31-47,115-216,122-189,154-168,179-203/Disulfide bonds: #status experi
F:158,60,183/Active site: His, Asp, Ser #status experimental
F:58,60,63,68/Binding site: calcium (Glu, Asn, Val, Glu) #status experimental
F:131-132/Cleavage site: Lys-Ser (autolytic) #status experimental

Query Match 26.2%; Score 394.5; DB 1; Length 229;
Best Local Similarity 37.4%; Pred. No. 7.2e-27;
Matches 91; Conservative 41; Mismatches 86; Indels 25; Gaps 8;

Qy 23 KIIGGFTTTEENQWFAAIYRRHGGSVTVVCGSLISPCWVLSATHCFIDYPKEDYIV 82
Db 6 KIVGGYTCGANTVEPYQVSL-----NSGYHFCGSLINSQWVLSAHCY-----KSGIQV 54
Qy 83 YLGRSLRNSNTQGMKEFEVENLILHKDYSADTLAHNDIALLKIRSGECRCAQPSRTIQT 142
Db 55 RLGEDNINNVGEGNQFISAKSIVHPSYNNTL--NNDIMLIKLS---AASINSRVAS 108
Qy 143 ICLPSMNDPQFGTSCITGFKENSTDYLYPEOLKMTVVKLISHRECQPHYVGSVTT 202
Db 109 ISLPT--SCASAGTQCLISGWNTKSGTSYPPVLYKCLKAPILSDSSCKSA--YPGDITS 164
Qy 203 KMLCAADPOWKTDSQGDGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLP 262
Db 165 NMFCAGYLEGGKDSQGDGGPVCNS--GK--LQGIWSWIGSGCAQKNGPVYTKVCNVVS 220
Qy 263 WIR 265
Db 221 WIK 223

RESULT 29
B61545
plasmin (EC 3.4.21.7) precursor - sheep (fragments)
N:Alternate names: plasminogen
N:Contains: miniplasminogen
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 28-Oct-1994 #sequence_revision 01-Nov-1996 #text_change 17-Mar-1999
C:Accession: B61545; S28200
R:Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A:Title: Structural aspects of the plasminogen of various species.
A:Reference number: A61545; MUID:89005015; PMID:3168975
A:Accession: B61545
A:Molecule type: protein
A:Residues: 1-37;38-117 <SCH>
R:Schaller, J.; Straub, C.; Kaempfer, U.; Rickli, E.E.
```



Protein Seq. Data Anal. 5, 21-25, 1992  
 A:Title: Complete amino acid sequence of ovine miniplasminogen.  
 A:Reference number: S28200; MUID:93149395; PMID:1492092  
 A:Accession: S28200  
 A:Molecule type: protein  
 A:Residues: 118-460 <SC2>  
 C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology  
 C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase; z  
 F:1-37,38-117,118-460/Product: plasminogen (fragments) #status experimental <PRO>  
 F:1-37/Domain: activation peptide (fragment) #status experimental <APT>  
 F:38-117,118-230,231-460/Product: plasmin (fragments) #status experimental <MAT>  
 F:41-118/Domain: kringle homology <KR4>  
 F:118-460/Product: miniplasminogen #status experimental <MIN>  
 F:132-211/Domain: kringle homology <KR5>  
 F:226-460/Domain: plasmin chain B #status experimental <BCH>  
 F:231-453/Domain: trypsin homology <TRY>  
 F:272,315,410/Active site: His, Asp, Ser #status predicted

Query Match 26.0%; Score 392.5; DB 2; Length 460;  
 Best Local Similarity 36.0%; Pred. No. 2.5e-26;  
 Matches 96; Conservative 35; Mismatches 107; Indels 29; Gaps 7;

QY 5 PPEELKFGCGQKTLRPR---FKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGGSLISP 61  
 DB 209 PQCESFDCGKPKVEPKCPAIVGCGVATPHSPWQVSLRRSR---EHFCGGTLISP 264  
 QY 62 CWVISATHCFIDYPPKKEDYIVVILGRSLNSNTQGENKFEVENLIHKDYSADTLAHNDI 121  
 DB 265 EWLTAHCLDILSGFSYTVILGAHYEMAREASVQETIPVSLFLEPSRA-----DI 316  
 QY 122 ALLKIRSGRCACQAPRTIOTCLPSMYNDPOF---GTSCEITGFGKENSIDYLYPQL 177  
 DB 317 ALLKLSPP---AVITDEVIPACLPs---PNYVADKIVCVITGGETOGT--FGVGR 366  
 QY 178 KMTVVKLISHRECQPHYGVSEVTTKCLCAADPQWKTDCSQGSGGPLVCSLQGRMTITG 237  
 DB 367 KEARLPVIENTKVCNRYEVLNGRVKSTELCAGLAGGTDSCQDSCGGLVCFEKRYYLQ 426  
 QY 238 IVSWGFGCALCKDKPGVYTVVSHFLPMI 264  
 DB 427 VTSWGLGCRPNKPGVYTVVSHFLPMI 453

RESULT 30  
 S13813  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 02-Dec-1993 #sequence\_revision 03-Aug-1995 #text\_change 22-Jun-1999  
 C:Accession: S13813  
 R:Le Huerou, I.; Wicker, C.; Guilloteau, P.; Toullec, R.; Puigserver, A.  
 Eur. J. Biochem. 193, 767-773, 1990  
 A:Title: Isolation and nucleotide sequence of cDNA clone for bovine pancreatic anionic  
 A:Reference number: S13813; MUID:91065383; PMID:1701147  
 A:Accession: S13813  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-247 <HUE>  
 A:Cross-references: EMBL:X54703; NID:9829; PIDN:CAA38513.1; PID:g830  
 C:Superfamily: trypsin; trypsin homology  
 C:Keywords: hydrolase; protein digestion; serine proteinase  
 F:24-239/Domain: trypsin homology <TRY>  
 F:63,107,200/Active site: His, Asp, Ser #status predicted

Query Match 25.9%; Score 390.5; DB 2; Length 247;  
 Best Local Similarity 38.1%; Pred. No. 1.8e-26;  
 Matches 93; Conservative 41; Mismatches 83; Indels 27; Gaps 9;

QY 23 KITGGEFTTIENQ-PFAAIYRRHGGSVTVVCGGSLISPCWISATHCFIDYPPKKEDYI 81  
 DB 23 KIVGG-YTCAENSPYQVSLNAGY-----HFCGSLINDQWVWSAAHY-----QYH 70  
 QY 82 VYLGRSLNSNTQGENKFEVENLIHKDYSADTLAHNDIALKIRSGRCACQAPRTIQ 141

DB 71 VRLGEYNIDVLEGGEOFDASKIIRHPKYSSWTL--DNDILLIKLSTP-----AVINARVS 124  
 QY 142 TICLPWMYNDPFGTSCETITGFKENSTDYLYPEOLKMTVVKLISHRECQPHYGVSEVT 201  
 DB 125 TILLPSAC--ASAGTECLISGKNTLSGWNYPDLQLCLVAPLISHADCEAS--YPCQIT 180  
 QY 202 TKMLCAADPQWKTDCSQGSGGPLVCSLQGRMTITGIVSWGRCALKDKGVYTVVSHFL 261  
 DB 181 NNMICAFLEGKGGKSCQSGSGGVACNQ----LOGIVSWGYCAOKGKPGVTVKVCNV 236  
 QY 262 PWIR 265  
 DB 237 DWIQ 240

RESULT 31  
 A21195  
 C:Species: Canis lupus familiaris (dog)  
 C:Date: 27-Jul-1990 #sequence\_revision 27-Jul-1990 #text\_change 22-Jun-1999  
 C:Accession: A21195  
 R:Pinsky, S.D.; LaForge, K.S.; Luc, V.; Scheele, G.  
 Proc. Natl. Acad. Sci. U.S.A. 80, 7486-7490, 1983  
 A:Title: Identification of cDNA clones encoding secretory isoenzyme forms: seq  
 A:Reference number: A21195; MUID:84170253; PMID:6584866  
 A:Accession: A21195  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-263 <GN>  
 A:Cross-references: GB:X01173; NID:g163945; PIDN:AAA30841.1; PID:g163946  
 C:Superfamily: trypsin; trypsin homology  
 C:Keywords: hydrolase; protein digestion; serine proteinase  
 F:34-256/Domain: trypsin homology <TRY>  
 F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 25.9%; Score 390; DB 2; Length 263;  
 Best Local Similarity 33.2%; Pred. No. 2.1e-26;  
 Matches 86; Conservative 51; Mismatches 100; Indels 22; Gaps 7

QY 11 FCGQKTLRPRF---KIIGGEFTTIENQWFAAIYRRHGGSVTVVCGGSLISPQWVIS 66  
 DB 17 FCGVPAIQVLSGLSRVINGEDAVPGSWQVSL----QDSTGFHFCGSLISEDWVVT 72  
 QY 67 ATHCFIDYPPKKEDYIVVILGRSLNSNTQGENKFEVENLIHKDYSADTLAHNDIALK 126  
 DB 73 AAHCGV---RTTHQVAGEFDQGSDAESIQLVKIAKVFKNPKFNMTI--NNDITLLKL 126  
 QY 127 RSKEGRCAQPSRTIOTCLPSMYNDPOFGTSCETITGFGKENSIDYLYPEOLKMTVVKLIS 186  
 DB 127 ATP-----ARSKTVSAVCLPQADDPFAGTLCVTTGELTKHTNANTPKLQQAALPLS 182  
 QY 187 HRECQPHYGVSEVTTKMLCAADPQWKTDCSQGSGGPLVCSLQGRMTITGIVSWGRC 246  
 DB 183 NAECKK--FWGSKITDLWVCAGAS--GVSSCGWSDSGGPLVCQKDGAWTLGVISWGS 238  
 QY 247 LKDKPGVYTVVSHFLPMIR 265  
 DB 239 STSTEGVYARVTKLIPWVQ 257

RESULT 32  
 S33777  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 18-Jun-1999  
 C:Accession: S33777; S32013  
 R:Farley, D.; Raymond, F.; Nick, H.  
 Biochim. Biophys. Acta 1173, 350-352, 1993  
 A:Title: Cloning and sequence analysis of rat hepsin, a cell surface serine pr  
 A:Reference number: S33777; MUID:93305733; PMID:8318546  
 A:Accession: S33777  
 A:Status: preliminary  
 A:Molecule type: mRNA

A; Gene: PK

```
F/110-199/Domain: apple repeat <Ap2>
F/200-289/Domain: apple repeat <Ap3>
F/291-380/Domain: apple repeat <Ap4>
```

```
F:391-621/Domain: trypsin homology <TRY>
F:21-104,47-77,51-57,111-194,137-166,141-147,201-284,227-256,231-237,292-375,334-73,
F:127,215,308,453,459,494/Binding site: carbohydrate (covalent) #status:
dicted
```

F:434,483,578/Active site: His, Asp, Ser #status predicted  
Query Match 25.5%; Score 385; DB 1; Length 638;

Matches 85; Conservative 56; Mismatches 106; Indels 16; Gaps 6

QY 17 TLRPRFKIIGBEFTTIENQFWFAAIYRRHRGGSVTYVCGGSLISPCWVISATHCFDYPK 76

DB 384 TTKINARIIVGGTNSLGEWPQVSL--QVKLVSQNHMCGGSIIGRQWILTTAAHCFDGIPI 441

QV 77 KEDYIWWLGRSLNSNTQGEKMFVENIILHKDYSADTLAHNDIALLKIRSKEGCAQP 136

DB 442 PDVRIYGGIINLSBITNKTFFSSIKELIIHQVKMSEGSY--DIALIKLQTP-----LNY 495

CV 137 SPTIOTICTPSMYNDPOPGTSCBITGFG---KENSTDYLYPEQLKMTVWKLISHRECCQP 193

Db 496 TEFQPCICLSKADNTIYTNCWITGWYKERTQNI----LQRATIPLVNEBQK- 550

DB 551 KYRDIYITKOMICAGYKEGGIDACKDGGGPLYCKHSGRWLVGITSWEGGCARKEQPGV 610

Db 611 YTKAEYIDWILEKIQSSKERAL 633

RESULT 34  
PLPG  
plasmin (EC 3.4.21.7) precursor - pig (fragment)

N:Contains: miniplasminogen  
C:Species: Sus scrofa domestica (domestic pig)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 07-Sep-1990 #sequence revision 01-Nov-1996 #text\_change 18-Jul-1997

R. Schaller, J.; Marti, T.; Roesselet, S.J.; Kaempfer, U.; Rickli, E.E. Fibrinolysis 1, 91-102, 1987

A:Accession: S03733  
A:Molecule type: protein  
A:residues: 1-550 <SCH>

Eur. J. Biochem. 114, 465-470, 1981  
 A:Title: Comparison of the primary structure of the N-terminal CNBR fragments  
 of human  
 A:Reference number. S01735: PMID:72318497  
 A:Reference number. S01735: PMID:72318497

A/Accession: 00357  
A/Molecule type: protein  
A/Residues: 1-57 <BRU>  
A/Residues: 1-57 <BRU>

A;Title: Determination of the complete amino-acid sequence of porcine miniplas: rogen

A:Reference number: A25834; MUID:65203907; PMID:3846533  
A:Accession: A25834  
A:Molecule type: protein  
A:Residues: 450-790 <VAR>  
C:Function:  
A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a variety of the graafian follicle; also activates the urokinase-type plasminogen activator  
A:Pathway: fibrinolysis  
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology  
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kidney; kringle; plasma; serine protease  
F:1-790/Product: plasminogen #status predicted <PRO>  
F:1-77/DNA: plasminogen-related protein precursor homology (fragment) <PLPH>  
F:1-77/DNA: activation peptide #status predicted <APT>  
F:78-560/Product: plasmin chain A #status predicted <ACH>  
F:166-243/DNA: kringle homology <KR1>  
F:166-243/DNA: kringle homology <KR2>  
F:256-333/DNA: kringle homology <KR3>  
F:358-435/DNA: kringle homology <KR4>  
F:450-790/Product: miniplasminogen #status experimental <MIN>  
F:461-540/DNA: kringle homology <KR5>  
F:561-790/Product: plasmin chain B #status experimental <BCH>  
F:561-793/DNA: trypsin homology <TRY>  
F:30-54,34-42,84-162,105-145,133-157,166-243,169-297,187-226,215-238,256-333,277-316,305-333,316-305  
bonds: #status predicted  
F:602,645,740/Active site: His, Asp, Ser #status predicted

Query Match 25.4%; Score 383.5; DB 1; Length 790;  
Best Local Similarity 36.0%; Pred. No. 2.8e-25;  
Matches 94; Conservative 39; Mismatches 99; Indels 29; Gaps 8;

QY 11 FQCGKTLRPF---FKIIGETTTIENQWFAAIYRHRGGSVTVVCGGSLISPCWVISA 67  
DB 545 FDGKPKVEKPKFARVGGCVIPHSWPMQISLRVYRG---HFCGGLTISPVELTA 600  
QY 68 THCFIDYPKKEDYIVVGLGRSLNNTQGMKFEVENLILHKDYSADTLAHHNDIALKLR 127  
DB 601 KHCLEKSSPSYKVIILGAEHYHGLGVEIDVSKLF--KEPS-----EADIALKL 652  
QY 128 SKEGRCAQPSRTIOTICLPSMNDPQF---GNSCEITGFGKENSVDLYPEQLKMTVVK 183  
DB 653 SP-----AVITDKVIPACLPF-----PNVAVDRACVITGWEYKGT--YGAGLLKEARLP 702  
QY 184 LISHRCQPHYGVSEVTTKMLCAADPQWKTDCQDGGSGPLVCSLQGRMTLTIGVSWGR 243  
DB 703 VIENKVCNRYELGGKVPNELCAGHLAGGIDSCQDGGSGPLVCFPEKDKVILQGVTSWGL 762  
QY 244 GCALKDKPGVYTRVSHPLPW 264  
DB 763 GCALPNKPGVYTRVSRVETW 783

RESULT 35  
KYRTE  
Chymotrypsin (EC 3.4.21.1) B precursor - rat  
N:Alternate names: chymotrypsinogen B  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 18-Jun-1999  
C:Accession: A22658  
J: Biol. Chem. 259, 14265-14270, 1984  
A:Title: Isolation and sequence of a rat chymotrypsin B gene.  
A:Reference number: A22658; MUID:85054881; PMID:6209274  
A:Accession: A22658  
A:Molecule type: DNA  
A:Residues: 1-263 <BEL>  
A:Cross-references: GB:K02298; NID:G203653; PIDN:AAA98732.1; PID:G203654  
C:Genetics:  
A:Introns: 18/1; 52/3; 79/2; 105/3; 166/1; 210/3  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase  
F:1-18/DNA: signal sequence #status predicted <SIG>  
F:19-33/DNA: propeptide #status predicted <PRO>  
F:34-263/Product: chymotrypsin B #status predicted <MAT>

Query Match 25.2%; Score 380; DB 2; Length 269;  
Best Local Similarity 34.1%; Pred. No. 1.6e-25;  
Matches 88; Conservative 50; Mismatches 106; Indels 14; Gaps 8

QY 13 CQCKTLRPF-KIIGETTTIENQWFAAIYRHRGGSVTVVCGGSLISPCWVISAHCF 71  
DB 17 CGLPANLPQLRVVGEDDARNSWPMQVSL--QYSSGQWRTCCGTLVDQSWVITAAHCI 75  
QY 72 IDPKKEDYIVVGLGRSLNNTQGMKFEVENLILHKDYSADTLAHHNDIALKIRKEG 131  
DB 76 ---SSRYRVVGLGRSLSTNEPGSLAVKSKLVVHQDWSNQLSNGNDIALKLASP-- 130  
QY 132 RQAPSRITOTICLPSMNDPQFSGTSCITGFGKENSVDLYPEQLKMTVVKLISHRECO 191  
DB 131 --VSLTDRKIQGLCPAAGTILPNVYCVYTWGR--LQTNASPDQLQCGQLLVVDYATCS 187  
QY 192 QPHYGVSEVTTKMLCAADPQWKTDCQDGGSGPLVCSLQGRMTLTIGVSWGR--GCALK 248  
DB 188 KEGWGSIVKTNMIGAGG-DGIISGDSGSGPLNCGQANGQVHGIVSFGSLGCVNY 246  
QY 249 DRPGVYTRVSHPLPW 266

F:34-256/DNA: trypsin homology <TRY>  
F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 25.3%; Score 382; DB 1; Length 263;  
Best Local Similarity 33.6%; Pred. No. 1e-25;  
Matches 87; Conservative 47; Mismatches 103; Indels 22; Gaps 7

QY 11 FQCGKTLRPF---KIIGETTTIENQWFAAIYRHRGGSVTVVCGGSLISPCWVISA 66  
DB 17 FQCGVPTIQVLTGLSRVNGEDAIFGSWPMQVSLQDK---TGFHFCGGLSIEDWVVT 72  
QY 67 ATHCFIDYPKKEDYIVVGLGRSLNNTQGMKFEVENLILHKDYSADTLAHHNDIALK 126  
DB 73 AARGV---KTSDDVVVAGEFDQGSDEENTQVLKIAQFKNPKFNFV--RNDITLLKL 126  
QY 127 RSKEGRCAQPSRTIOTICLPSMNDPQFSGTSCITGFGKENSVDLYPEQLKMTVVK 186  
DB 127 ATP---AQFSETVSAVCLNEDDFPPGTVCATCTGKTKYNALKTPEKLOQAALFIS 182  
QY 187 HRECOQPHYGVSEVTTKMLCAADPQWKTDCQDGGSGPLVCSLQGRMTLTIGVSWGRGA 246  
DB 183 EADCKKS--WGSKITDVTMTAGAS--GVSSCGDGGSGPLVCGQGWTLTLAGIVSWSGVC 238  
QY 247 LKDRPGVYTRVSHPLPW 265  
DB 239 STSTPAVYTRVSHPLPW 257

RESULT 36  
A26823  
pancreatic elastase II (EC 3.4.21.71) precursor - pig  
N:Alternate names: pancreatoelastidase E  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 16-Aug-1988 #sequence\_revision 16-Aug-1988 #text\_change 22-Jun-1999  
C:Accession: A26823  
R:Kawashima, I.; Tani, T.; Shimoda, K.; Takiguchi, Y.  
DNA 6, 163-172, 1987  
A:Title: Characterization of pancreatic elastase II cDNAs: two elastase II mRNAs are encoded by two genes.  
A:Reference number: A20958; MUID:87217962; PMID:3646943  
A:Accession: A26823  
A:Molecule type: mRNA  
A:Residues: 1-269 <KAN>  
A:Cross-references: GB:M16651; NID:G164441; PIDN:AAA31027.1; PID:G164442  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; serine proteinase  
F:1-16/DNA: signal sequence #status predicted <SIG>  
F:17-28/DNA: propeptide #status predicted <PRO>  
F:29-269/Product: elastase II #status predicted <MAT>  
F:29-262/DNA: trypsin homology <TRY>  
F:73,121,216/Active site: His, Asp, Ser #status predicted

Db 247 HKPSVTRVSNVIDWINS 264

# RESULT 37

S33496  
 tryptsin (BC 3.4.21.4) IV form a - human  
 C:Species: Homo sapiens (man)  
 C>Date: 03-Mar-1994 #sequence\_revision 03-Aug-1995 #text\_change 15-Aug-1997  
 C:Accession: S33496  
 R:Wiegand, U.; Corbach, S.; Minn, A.; Kang, J.; Mueller-Hill, B.  
 Submitted to the EMBL Data Library, March 1993  
 A:Description: Identification, cloning and characterization of a cDNA encoding a human b  
 A:Reference number: S33496  
 A:Accession: S33496  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-304 <WIE>  
 A:Cross-references: EMBL:X72781  
 C:Genetics:  
 A:Gene: GDB:PRSS4; TRY4  
 A:Cross-references: GDB:335300  
 A:Map position: 7q35-7q35  
 C:Superfamily: trypsin; trypsin homology  
 C:Keywords: hydrolase; serine proteinase  
 F:81-296/Domain: trypsin homology <TRY>  
 F:120,164,257/Active site: His, Asp, Ser #status predicted

Query Match 25.2%; Score 379.5; DB 2; Length 304;  
 Best Local Similarity 37.7%; Pred. No. 2e-25;  
 Matches 92; Conservative 38; Mismatches 87; Indels 27; Gaps 9;  
 QY 23 KIIGGEFTTIENQ-PWFAAIYRRHGGSVTVYVCGSLISPCWVISATHCFIDYPKKEDIY 81  
 DB 80 KIVGG-YTCENSLPYQVSL-----NSGSHFCGSLISEQWVWSAAHCY-----KTRIQ 127  
 QY 82 VYLGSRSLNNTQGEKKFEVENILHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQ 141  
 DB 128 VRLGEHNKIVLENGEFINAAKILRHPKYNRDL--DNDIMLIKSSP----AVINARVS 181  
 QY 142 TICLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMTVVKLIISHRECCQPHYGSEVT 201  
 DB 182 TISLPTA--PPAAGTECLISGWNTLSFGADYDELKCLDAPVLTOAECKAS--YPGKIT 237  
 QY 202 TKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVSGRCALKDKPGVYTRVSHFL 261  
 DB 238 NSMFCVGFLEGGKDSQORDSGGPFVVCNGQ-----LQGVWSHGHCWKNRPGVYTRVY 293  
 QY 262 PWIR 265  
 DB 294 DWIK 297

RESULT 38  
 I38136  
 chymotrypsin-like proteinase (BC 3.4.21.-) CTRL-1 - human  
 C:Species: Homo sapiens (man)  
 C>Date: 06-Sep-1996 #sequence\_revision 06-Sep-1996 #text\_change 22-Jun-1999  
 C:Accession: I38136  
 R:Larsen, F.; Solheim, J.; Kristensen, T.; Kolsto, A.B.; Prydz, H.  
 Hum. Mol. Genet. 2, 1589-1595, 1993  
 A:Title: A tight cluster of five unrelated human genes on chromosome 16q22.1.  
 A:Reference number: I38135; MUID:94093544; PMID:8268911  
 A:Accession: I38136  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-264 <RES>  
 A:Cross-references: EMBL:X71874; NID:g406226; PIDN:CAA50710.1; PID:g406228  
 C:Genetics:  
 A:Gene: GDB:CTRL  
 A:Map position: 16q22.1-16q22.1  
 A:Cross-references: GDB:204061  
 A:Introns: 18/1; 52/3; 79/2; 106/3; 167/1; 211/3  
 C:Superfamily: trypsin; trypsin homology

C:Keywords: hydrolase; serine proteinase  
 F:34-257/Domain: trypsin homology <TRY>  
 F:75,121,214/Active site: His, Asp, Ser #status predicted

Query Match 25.1%; Score 378.5; DB 2; Length 264;  
 Best Local Similarity 35.0%; Pred. No. 2.1e-25;  
 Matches 90; Conservative 43; Mismatches 101; Indels 23; Gaps 10;  
 QY 13 CGQKTLRPRP-----KIIGGEFTTIENQ-PWFAAIYRRHGGSVTVYVCGSLISPCWISAT 68  
 DB 19 CGIPAKPALSPQRIQVNGENAVLGSPWQVSL-----QDSSGFHFCGSLISQSWVTAA 74  
 QY 69 HCFIDYPKKEDIYVYLGSRSLNNTQGEKKFEVENILHKDYSADTLAHNDIALLKIRS 128  
 DB 75 HCNVS-PGR--HFVVLGEYDRSSNAEPLOVLVSRAITHPSWNSITM--NNDVTLLKLAS 129  
 QY 129 KEGRCAQPSRTITQICLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMTVVKLIISHR 186  
 DB 130 P-----AQYTTIRISFVCLASSNEALTEGLTCVTTGWGRLSGVGNVTFPAHQQVALPLVTYN 185  
 QY 189 ECQPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVSWG-RCAL 247  
 DB 186 QCRQ--YWGSSITDSMICAGGA--GASSCQDGGPLVCQKGMTWTWLVIGVSWGTKNCNV 241  
 QY 248 KDKPGVYTRVSHFLPMI 264  
 DB 242 R-APAVYTRVSKFTWI 257

# RESULT 39

I38363  
 trypsin (BC 3.4.21.4) IV form b precursor - human (fragment)  
 C:Species: Homo sapiens (man)  
 C>Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 23-Aug-1997  
 C:Accession: I38363; S37538  
 R:Wiegand, U.; Corbach, S.; Minn, A.; Kang, J.; Mueller-Hill, B.  
 Gene 136, 167-175, 1993  
 A:Title: Cloning of the cDNA encoding human brain trypsinogen and characteriza  
 A:Reference number: I38363; MUID:94123994; PMID:8294000  
 A:Accession: I38363  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-259 <RES>  
 A:Cross-references: EMBL:X71345; NID:g405755  
 C:Genetics:  
 A:Introns: 25/3; 78/1; 162/3; 208/2  
 C:Superfamily: trypsin; trypsin homology  
 C:Keywords: hydrolase; serine proteinase  
 F:37-251/Domain: trypsin homology <TRY>  
 F:75,119,212/Active site: His, Asp, Ser #status predicted

Query Match 25.1%; Score 378; DB 2; Length 259;  
 Best Local Similarity 37.0%; Pred. No. 2.3e-25;  
 Matches 90; Conservative 39; Mismatches 88; Indels 26; Gaps 8  
 QY 23 KIIGGEFTTIENQ-PWFAAIYRRHGGSVTVYVCGSLISPCWVISATHCFIDYPKKEDIY 82  
 DB 36 KIVGG-YTCENSLPYQVSL-----NSGSHFCGSLISEQWVWSAAHCY-----KTRIQ 83  
 QY 83 VYLGSRSLNNTQGEKKFEVENILHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQ 142  
 DB 84 RLGEHNKIVLENGEFINAAKILRHPKYNRDL--DNDIMLIKSSP----AVINARVST 137  
 QY 143 ICPLSMYNDPQGTSCITGFGKENSTDYLYPEQLKMTVVKLIISHRECCQPHYGSEVTT 202  
 DB 138 ISLPTA--PPAAGTECLISGWNTLSFGADYDELKCLDAPVLTOAECKAS--YPGKITN 193  
 QY 203 KMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLP 262  
 DB 194 SMFCVGFLEGGKDSQORDSGGPFVVCNGQ-----LQGVWSHGHCWKNRPGVYTRVY 249  
 QY 263 WIR 265



A,Molecule type: protein  
A,Residues: 22-27 <LI2>  
C,Comment: Plasminogen is synthesized by the kidney and is present in plasma and many of medially after dissociation from the clot. In the presence of the inhibitor, the active inhibitor the activation involves also removal of the activation peptide.  
C,Comment: Stromelysin 1 (see PIR:KCMSS1) acts on plasminogen to produce angiotatin. To be useful in treating solid tumors.

C,Function:  
A,Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a variety of the walls of the graafian follicle; also activates the urokinase-type plasminogen activator.  
A,Superfamily: fibrinolysis  
C,Pathway: plasmin; kringles; plasminogen-related protein precursor homology  
C,Keywords: angiotensin inhibitor; blood; duplication; fibrinolysis; glycoprotein; hydrolysis; plasminogen-related protein precursor homology  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-812/Product: plasminogen #status predicted <PRO>  
F:79-466/Product: angiotatin #status predicted <AST>  
F:97-581-812/Product: plasmin #status predicted <MAT>  
F:97-581/Domain: chain A #status predicted <ACH>  
F:103-181/Domain: kringles homology <KR1>  
F:185-262/Domain: kringles homology <KR2>  
F:275-352/Domain: kringles homology <KR3>  
F:377-454/Domain: kringles homology <KR4>  
F:481-560/Domain: kringles homology <KR5>  
F:582-812/Domain: chain B #status predicted <BCH>  
F:582-805/Domain: trypsin homology <TRY>  
F:49-73,53-61,103-124,124-164,152-176,185-262,188-316,206-245,234-257,275-352,296-335,324-335,336-337,338-339,340-341,342-343,344-345,346-347,348-349,350-351,352-353,354-355,356-357,358-359,360-361,362-363,364-365,366-367,368-369,370-371,372-373,374-375,376-377,378-379,380-381,382-383,384-385,386-387,388-389,390-391,392-393,394-395,396-397,398-399,400-401,402-403,404-405,406-407,408-409,410-411,412-413,414-415,416-417,418-419,420-421,422-423,424-425,426-427,428-429,430-431,432-433,434-435,436-437,438-439,440-441,442-443,444-445,446-447,448-449,450-451,452-453,454-455,456-457,458-459,460-461,462-463,464-465,466-467,468-469,470-471,472-473,474-475,476-477,478-479,480-481,482-483,484-485,486-487,488-489,490-491,492-493,494-495,496-497,498-499,500-501,502-503,504-505,506-507,508-509,510-511,512-513,514-515,516-517,518-519,520-521,522-523,524-525,526-527,528-529,530-531,532-533,534-535,536-537,538-539,540-541,542-543,544-545,546-547,548-549,550-551,552-553,554-555,556-557,558-559,560-561,562-563,564-565,566-567,568-569,570-571,572-573,574-575,576-577,578-579,580-581,582-583,584-585,586-587,588-589,590-591,592-593,594-595,596-597,598-599,600-601,602-603,604-605,606-607,608-609,610-611,612-613,614-615,616-617,618-619,620-621,622-623,624-625,626-627,628-629,630-631,632-633,634-635,636-637,638-639,640-641,642-643,644-645,646-647,648-649,650-651,652-653,654-655,656-657,658-659,660-661,662-663,664-665,666-667,668-669,670-671,672-673,674-675,676-677,678-679,680-681,682-683,684-685,686-687,688-689,690-691,692-693,694-695,696-697,698-699,700-701,702-703,704-705,706-707,708-709,710-711,712-713,714-715,716-717,718-719,720-721,722-723,724-725,726-727,728-729,730-731,732-733,734-735,736-737,738-739,740-741,742-743,744-745,746-747,748-749,750-751,752-753,754-755,756-757,758-759,760-761,762-763,764-765,766-767,768-769,770-771,772-773,774-775,776-777,778-779,780-781,782-783,784-785,786-787,788-789,790-791,792-793,794-795,796-797,798-799,800-801,802-803,804-805,806-807,808-809,810-811,812-813,814-815,816-817,818-819,820-821,822-823,824-825,826-827,828-829,830-831,832-833,834-835,836-837,838-839,840-841,842-843,844-845,846-847,848-849,850-851,852-853,854-855,856-857,858-859,860-861,862-863,864-865,866-867,868-869,870-871,872-873,874-875,876-877,878-879,880-881,882-883,884-885,886-887,888-889,890-891,892-893,894-895,896-897,898-899,900-901,902-903,904-905,906-907,908-909,910-911,912-913,914-915,916-917,918-919,920-921,922-923,924-925,926-927,928-929,930-931,932-933,934-935,936-937,938-939,940-941,942-943,944-945,946-947,948-949,950-951,952-953,954-955,956-957,958-959,960-961,962-963,964-965,966-967,968-969,970-971,972-973,974-975,976-977,978-979,980-981,982-983,984-985,986-987,988-989,990-991,992-993,994-995,996-997,998-999,1000-1001,1002-1003,1004-1005,1006-1007,1008-1009,1010-1011,1012-1013,1014-1015,1016-1017,1018-1019,1020-1021,1022-1023,1024-1025,1026-1027,1028-1029,1030-1031,1032-1033,1034-1035,1036-1037,1038-1039,1040-1041,1042-1043,1044-1045,1046-1047,1048-1049,1050-1051,1052-1053,1054-1055,1056-1057,1058-1059,1060-1061,1062-1063,1064-1065,1066-1067,1068-1069,1070-1071,1072-1073,1074-1075,1076-1077,1078-1079,1080-1081,1082-1083,1084-1085,1086-1087,1088-1089,1090-1091,1092-1093,1094-1095,1096-1097,1098-1099,1100-1101,1102-1103,1104-1105,1106-1107,1108-1109,1110-1111,1112-1113,1114-1115,1116-1117,1118-1119,1120-1121,1122-1123,1124-1125,1126-1127,1128-1129,1130-1131,1132-1133,1134-1135,1136-1137,1138-1139,1140-1141,1142-1143,1144-1145,1146-1147,1148-1149,1150-1151,1152-1153,1154-1155,1156-1157,1158-1159,1160-1161,1162-1163,1164-1165,1166-1167,1168-1169,1170-1171,1172-1173,1174-1175,1176-1177,1178-1179,1180-1181,1182-1183,1184-1185,1186-1187,1188-1189,1190-1191,1192-1193,1194-1195,1196-1197,1198-1199,1200-1201,1202-1203,1204-1205,1206-1207,1208-1209,1210-1211,1212-1213,1214-1215,1216-1217,1218-1219,1220-1221,1222-1223,1224-1225,1226-1227,1228-1229,1230-1231,1232-1233,1234-1235,1236-1237,1238-1239,1240-1241,1242-1243,1244-1245,1246-1247,1248-1249,1250-1251,1252-1253,1254-1255,1256-1257,1258-1259,1260-1261,1262-1263,1264-1265,1266-1267,1268-1269,1270-1271,1272-1273,1274-1275,1276-1277,1278-1279,1280-1281,1282-1283,1284-1285,1286-1287,1288-1289,1290-1291,1292-1293,1294-1295,1296-1297,1298-1299,1300-1301,1302-1303,1304-1305,1306-1307,1308-1309,1310-1311,1312-1313,1314-1315,1316-1317,1318-1319,1320-1321,1322-1323,1324-1325,1326-1327,1328-1329,1330-1331,1332-1333,1334-1335,1336-1337,1338-1339,1340-1341,1342-1343,1344-1345,1346-1347,1348-1349,1350-1351,1352-1353,1354-1355,1356-1357,1358-1359,1360-1361,1362-1363,1364-1365,1366-1367,1368-1369,1370-1371,1372-1373,1374-1375,1376-1377,1378-1379,1380-1381,1382-1383,1384-1385,1386-1387,1388-1389,1390-1391,1392-1393,1394-1395,1396-1397,1398-1399,1400-1401,1402-1403,1404-1405,1406-1407,1408-1409,1410-1411,1412-1413,1414-1415,1416-1417,1418-1419,1420-1421,1422-1423,1424-1425,1426-1427,1428-1429,1430-1431,1432-1433,1434-1435,1436-1437,1438-1439,1440-1441,1442-1443,1444-1445,1446-1447,1448-1449,1450-1451,1452-1453,1454-1455,1456-1457,1458-1459,1460-1461,1462-1463,1464-1465,1466-1467,1468-1469,1470-1471,1472-1473,1474-1475,1476-1477,1478-1479,1480-1481,1482-1483,1484-1485,1486-1487,1488-1489,1490-1491,1492-1493,1494-1495,1496-1497,1498-1499,1500-1501,1502-1503,1504-1505,1506-1507,1508-1509,1510-1511,1512-1513,1514-1515,1516-1517,1518-1519,1520-1521,1522-1523,1524-1525,1526-1527,1528-1529,1530-1531,1532-1533,1534-1535,1536-1537,1538-1539,1540-1541,1542-1543,1544-1545,1546-1547,1548-1549,1550-1551,1552-1553,1554-1555,1556-1557,1558-1559,1560-1561,1562-1563,1564-1565,1566-1567,1568-1569,1570-1571,1572-1573,1574-1575,1576-1577,1578-1579,1580-1581,1582-1583,1584-1585,1586-1587,1588-1589,1590-1591,1592-1593,1594-1595,1596-1597,1598-1599,1600-1601,1602-1603,1604-1605,1606-1607,1608-1609,1610-1611,1612-1613,1614-1615,1616-1617,1618-1619,1620-1621,1622-1623,1624-1625,1626-1627,1628-1629,1630-1631,1632-1633,1634-1635,1636-1637,1638-1639,1640-1641,1642-1643,1644-1645,1646-1647,1648-1649,1650-1651,1652-1653,1654-1655,1656-1657,1658-1659,1660-1661,1662-1663,1664-1665,1666-1667,1668-1669,1670-1671,1672-1673,1674-1675,1676-1677,1678-1679,1680-1681,1682-1683,1684-1685,1686-1687,1688-1689,1690-1691,1692-1693,1694-1695,1696-1697,1698-1699,1700-1701,1702-1703,1704-1705,1706-1707,1708-1709,1710-1711,1712-1713,1714-1715,1716-1717,1718-1719,1720-1721,1722-1723,1724-1725,1726-1727,1728-1729,1730-1731,1732-1733,1734-1735,1736-1737,1738-1739,1740-1741,1742-1743,1744-1745,1746-1747,1748-1749,1750-1751,1752-1753,1754-1755,1756-1757,1758-1759,1760-1761,1762-1763,1764-1765,1766-1767,1768-1769,1770-1771,1772-1773,1774-1775,1776-1777,1778-1779,1780-1781,1782-1783,1784-1785,1786-1787,1788-1789,1790-1791,1792-1793,1794-1795,1796-1797,1798-1799,1800-1801,1802-1803,1804-1805,1806-1807,1808-1809,1810-1811,1812-1813,1814-1815,1816-1817,1818-1819,1820-1821,1822-1823,1824-1825,1826-1827,1828-1829,1830-1831,1832-1833,1834-1835,1836-1837,1838-1839,1840-1841,1842-1843,1844-1845,1846-1847,1848-1849,1850-1851,1852-1853,1854-1855,1856-1857,1858-1859,1860-1861,1862-1863,1864-1865,1866-1867,1868-1869,1870-1871,1872-1873,1874-1875,1876-1877,1878-1879,1880-1881,1882-1883,1884-1885,1886-1887,1888-1889,1890-1891,1892-1893,1894-1895,1896-1897,1898-1899,1900-1901,1902-1903,1904-1905,1906-1907,1908-1909,1910-1911,1912-1913,1914-1915,1916-1917,1918-1919,1920-1921,1922-1923,1924-1925,1926-1927,1928-1929,1930-1931,1932-1933,1934-1935,1936-1937,1938-1939,1940-1941,1942-1943,1944-1945,1946-1947,1948-1949,1950-1951,1952-1953,1954-1955,1956-1957,1958-1959,1960-1961,1962-1963,1964-1965,1966-1967,1968-1969,1970-1971,1972-1973,1974-1975,1976-1977,1978-1979,1980-1981,1982-1983,1984-1985,1986-1987,1988-1989,1990-1991,1992-1993,1994-1995,1996-1997,1998-1999,2000-2001,2002-2003,2004-2005,2006-2007,2008-2009,2010-2011,2012-2013,2014-2015,2016-2017,2018-2019,2020-2021,2022-2023,2024-2025,2026-2027,2028-2029,2030-2031,2032-2033,2034-2035,2036-2037,2038-2039,2040-2041,2042-2043,2044-2045,2046-2047,2048-2049,2050-2051,2052-2053,2054-2055,2056-2057,2058-2059,2060-2061,2062-2063,2064-2065,2066-2067,2068-2069,2070-2071,2072-2073,2074-2075,2076-2077,2078-2079,2080-2081,2082-2083,2084-2085,2086-2087,2088-2089,2090-2091,2092-2093,2094-2095,2096-2097,2098-2099,2100-2101,2102-2103,2104-2105,2106-2107,2108-2109,2110-2111,2112-2113,2114-2115,2116-2117,2118-2119,2120-2121,2122-2123,2124-2125,2126-2127,2128-2129,2130-2131,2132-2133,2134-2135,2136-2137,2138-2139,2140-2141,2142-2143,2144-2145,2146-2147,2148-2149,2150-2151,2152-2153,2154-2155,2156-2157,2158-2159,2160-2161,2162-2163,2164-2165,2166-2167,2168-2169,2170-2171,2172-2173,2174-2175,2176-2177,2178-2179,2180-2181,2182-2183,2184-2185,2186-2187,2188-2189,2190-2191,2192-2193,2194-2195,2196-2197,2198-2199,2200-2201,2202-2203,2204-2205,2206-2207,2208-2209,2210-2211,2212-2213,2214-2215,2216-2217,2218-2219,2220-2221,2222-2223,2224-2225,2226-2227,2228-2229,2230-2231,2232-2233,2234-2235,2236-2237,2238-2239,2240-2241,2242-2243,2244-2245,2246-2247,2248-2249,2250-2251,2252-2253,2254-2255,2256-2257,2258-2259,2260-2261,2262-2263,2264-2265,2266-2267,2268-2269,2270-2271,2272-2273,2274-2275,2276-2277,2278-2279,2280-2281,2282-2283,2284-2285,2286-2287,2288-2289,2290-2291,2292-2293,2294-2295,2296-2297,2298-2299,2300-2301,2302-2303,2304-2305,2306-2307,2308-2309,2310-2311,2312-2313,2314-2315,2316-2317,2318-2319,2320-2321,2322-2323,2324-2325,2326-2327,2328-2329,2330-2331,2332-2333,2334-2335,2336-2337,2338-2339,2340-2341,2342-2343,2344-2345,2346-2347,2348-2349,2350-2351,2352-2353,2354-2355,2356-2357,2358-2359,2360-2361,2362-2363,2364-2365,2366-2367,2368-2369,2370-2371,2372-2373,2374-2375,2376-2377,2378-2379,2380-2381,2382-2383,2384-2385,2386-2387,2388-2389,2390-2391,2392-2393,2394-2395,2396-2397,2398-2399,2400-2401,2402-2403,2404-2405,2406-2407,2408-2409,2410-2411,2412-2413,2414-2415,2416-2417,2418-2419,2420-2421,2422-2423,2424-2425,2426-2427,2428-2429,2430-2431,2432-2433,2434-2435,2436-2437,2438-2439,2440-2441,2442-2443,2444-2445,2446-2447,2448-2449,2450-2451,2452-2453,2454-2455,2456-2457,2458-2459,2460-2461,2462-2463,2464-2465,2466-2467,2468-2469,2470-2471,2472-2473,2474-2475,2476-2477,2478-2479,2480-2481,2482-2483,2484-2485,2486-2487,2488-2489,2490-2491,2492-2493,2494-2495,2496-2497,2498-2499,2500-2501,2502-2503,2504-2505,2506-2507,2508-2509,2510-2511,2512-2513,2514-2515,2516-2517,2518-2519,2520-2521,2522-2523,2524-2525,2526-2527,2528-2529,2530-2531,2532-2533,2534-2535,2536-2537,2538-2539,2540-2541,2542-2543,2544-2545,2546-2547,2548-2549,2550-2551,2552-2553,2554-2555,2556-2557,2558-2559,2560-2561,2562-2563,2564-2565,2566-2567,2568-2569,2570-2571,2572-2573,2574-2575,2576-2577,2578-2579,2580-2581,2582-2583,2584-2585,2586-2587,2588-2589,2590-2591,2592-2593,2594-2595,2596-2597,2598-2599,2600-2601,2602-2603,2604-2605,2606-2607,2608-2609,2610-2611,2612-2613,2614-2615,2616-2617,2618-2619,2620-2621,2622-2623,2624-2625,2626-2627,2628-2629,2630-2631,2632-2633,2634-2635,2636-2637,2638-2639,2640-2641,2642-2643,2644-2645,2646-2647,2648-2649,2650-2651,2652-2653,2654-2655,2656-2657,2658-2659,2660-2661,2662-2663,2664-2665,2666-2667,2668-2669,2670-2671,2672-2673,2674-2675,2676-2677,2678-2679,2680-2681,2682-2683,2684-2685,2686-2687,2688-2689,2690-2691,2692-2693,2694-2695,2696-2697,2698-2699,2700-2701,2702-2703,2704-2705,2706-2707,2708-2709,2710-2711,2712-2713,2714-2715,2716-2717,2718-2719,2720-2721,2722-2723,2724-2725,2726-2727,2728-2729,2730-2731,2732-2733,2734-2735,2736-2737,2738-2739,2740-2741,2742-2743,2744-2745,2746-2747,2748-2749,2750-2751,2752-2753,2754-2755,2756-2757,2758-2759,2760-2761,2762-2763,2764-2765,2766-2767,2768-2769,2770-2771,2772-2773,2774-2775,2776-2777,2778-2779,2780-2781,2

```

Db 125 TISLPRACAP--GTQCLISGNGTSLSSGTNYPELLQCLDAPILOTAQCEAS--YPGQIT 180
QY 202 TWMLCAADPQWKTDSCQDSGGPLVCSLQGRMTLTGIVSWGRCALXDKPGVYTRVGHFL 261
Db 181 ENWICAGFLEGKDSQDSGGPVCNGB-----LQIVSWGRCALXDKPGVYTRVGHFL 236
QY 262 PWIRS 266
Db 237 DWIQS 241

RESULT 45
A25606
Tissue kallikrein (EC 3.4.21.35) submandibular precursor - mouse
N:Alternate names: glandular kallikrein; kininogenin
C:Species: Mus musculus (house mouse)
C:Date: 19-Nov-1988 #sequence_revision 30-Jun-1991 #text_change 13-Nov-1998
C:Accession: A25606; S06661
R:van Leeuwen, B.H.; Evans, B.A.; Tregear, G.W.; Richards, R.I.
J. Biol. Chem. 261, 5529-5535, 1986
A:Title: Mouse glandular kallikrein genes. Identification, structure, and expression of
A:Reference number: A25606; MUID:86168299; PMID:3007510
A:Accession: A25606
A:Molecule type: DNA
A:Residues: 1-261 <VAN>
A:Cross-references: GB:M13498
A:Experimental source: renal
A>Note: The authors translated the codon GTC for residue 57 as Ala, CTC for residue 105
R:Murakami, K.; Ikigai, H.; Nagumo, N.; Tomita, M.; Shimamura, T.
FEBS Lett. 257, 400-402, 1989
A:Title: A cytosolic tissue kallikrein isolated from mouse submandibular glands.
A:Reference number: S06660; MUID:90060365; PMID:2583286
A:Accession: S06661
A:Molecule type: protein
A:Residues: 165-174 <MUR>
C:Genetics:
A:Gene: MGK-6
A:Introns: 16/1; 69/2; 165/1; 210/3
C:Superfamily: trypsin; trypsin homology
C:Keywords: hydrolase; serine proteinase
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-261/Product: tissue kallikrein #status predicted <WAT>
F:25-253/Domain: trypsin homology <TRY>
F:65,120,213/Active site: His, Asp, Ser #status predicted

Query Match 24.8%; Score 373.5; DB 2; Length 261;
Best Local Similarity 33.8%; Pred. No. 5.7e-25;
Matches 88; Conservative 43; Mismatches 96; Indels 33; Gaps 9;

QY 23 KIIGSEFTTIENQFWFAIYRRHGGSVTVVCGSLISPCWVISAHCFIDYPKEDYIV 92
Db 24 RIVGFGNCEKNSQPMQVAVYR-----FTKYQCGILLNVNVLTAACHND-----KYQV 73

QY 83 YLGRSLNNTQGMKFEVENLILHKDYSADTLAH-----NDIALKIRSKEGRC 133
Db 74 WLGNKFLDEPQAQRHLVSKALPHDFNLSLNEHTPQDPDYNDMLRLK-----KP 129

QY 134 AQPSTIQTICLPSMYNDPQFTSCETITGFGKNSDYLYPEQLKMTVVKLISHRECOQP 193
Db 130 ADITDVVKPDLPT--EELKLGSTCLASGWSITPVKYEYDPELQVNLKLLPNEDCAKA 187

QY 194 HYGSEVTTKMLCAADPQWKTDSCQDSGGPLVCSLQGRMTLTGIVSWGRC--CALXDKPG 252
Db 188 HI--EKVTDMLCAGMDGKDTCAAGSGGPLICD-----GVLQGITSWGTPCGKPNVPG 241

QY 253 VYTRVSHFLPWIRSHKTEEN 272
Db 242 IYTRVLNFTWIR--ETWAEN 260

RESULT 46
NGMSG

```

```

7S nerve growth factor gamma chain (EC 3.4.21.-) precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 18-Dec-1981 #sequence_revision 17-May-1985 #text_change 18-Jun-1999
C:Accession: A91005; A90949; A93510; A92341; A00942; A21093; A22705
R:Evans, B.A.; Richards, R.I.
EMBO J. 4, 133-138, 1985
A:Title: Genes for the alpha and gamma subunits of mouse nerve growth factor a
A:Reference number: A91005; MUID:85257431; PMID:3848399
A:Accession: A91005
A:Molecule type: DNA
A:Residues: 1-261 <EVA>
R:Ullrich, A.; Gray, A.; Wood, W.I.; Hayflick, J.; Seeburg, P.H.
DNA 3, 387-392, 1984
A:Title: Isolation of a cDNA clone coding for the gamma-subunit of mouse nerve
A:Reference number: A90949; MUID:85076169; PMID:6548955
A:Accession: A90949
A:Molecule type: mRNA
A:Residues: 1-261 <ULL>
A:Cross-references: GB:X01389; NID:953373; PIDN:CAA25645.1; PID:953374
R:Howles, P.N.; Dickinson, D.P.; DiCaprio, L.L.; Woodworth-Gutai, M.; Gross, K.
Nucleic Acids Res. 12, 2791-2805, 1984
A:Title: Use of a cDNA recombinant for the gamma-subunit of mouse nerve growth
A:Reference number: A93510; MUID:84169573; PMID:6200835
A:Accession: A93510
A:Molecule type: mRNA
A:Residues: 127-202, 'E', 204-261 <HOW>
A:Cross-references: GB:X00472; NID:954260; PIDN:CAA25154.1; PID:954261
A:Experimental source: inbred strain DBA/2J
R:Thomas, K.A.; Baglan, N.C.; Bradshaw, R.A.
J. Biol. Chem. 256, 9156-9166, 1981
A:Title: The amino acid sequence of the gamma-subunit of mouse submaxillary gl
A:Reference number: A92341; MUID:81264363; PMID:7263706
A:Accession: A92341
A:Molecule type: protein
A:Residues: 25-107, 112-261 <THO>
A:Experimental source: outbred strain Swiss Webster
C:Comment: 7S nerve growth factor is composed of two alpha chains, a beta dime
C:Genetics:
A:Map position: 7
A:Introns: 16/1; 69/2; 165/1; 210/3
C:Superfamily: trypsin; trypsin homology
C:Keywords: glycoprotein; growth factor; hydrolase; serine proteinase; submand
F:1-18/Domain: signal sequence #status predicted <SIG>
F:25-107/Domain: trypsin homology <TRY>
F:25-107/Domain: segment B1 <GB1>
F:112-107, 112-261/Product: nerve growth factor gamma chain (active form) #statu
F:112-261/Domain: segment A <GAA>
F:112-154/Domain: segment C <GCC>
F:165-261/Domain: segment B2 <GB2>
F:31-173, 50-66, 152-219, 184-198, 203-234/Disulfide bonds: #status predicted
F:65,120,213/Active site: His, Asp, Ser #status predicted
F:102/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 24.7%; Score 372; DB 1; Length 261;
Best Local Similarity 32.4%; Pred. No. 7.6e-25;
Matches 82; Conservative 48; Mismatches 91; Indels 32; Gaps 8

QY 23 KIIGSEFTTIENQFWFAIYRRHGGSVTVVCGSLISPCWVISAHCFIDYPKEDYIV 82
Db 24 RIVGFGNCEKNSQPMQVAVYR-----FTKYQCGILLNVNVLTAACHND-----NYKV 73

QY 83 YLGRSLNNTQGMKFEVENLILHKDYSADTLAH-----NDIALKIRSKEGRC 133
Db 74 WLGNKFLDEPQAQRHLVSKALPHDFNLSLNEHTPQDPDYNDMLRLK-----SKP 129

QY 134 AQPSTIQTICLPSMYNDPQFTSCETITGFGKNSDYLYPEQLKMTVVKLISHRECOQP 193
Db 130 ADITDVVKPDLPT--EELKLGSTCLASGWSITPVKYEYDPELQVNLKLLPNEDCAKA 187

QY 194 HYGSEVTTKMLCAADPQWKTDSCQDSGGPLVCSLQGRMTLTGIVSWGRC--GCALXDKPG 252
Db 188 HI--EKVTDMLCAGMDGKDTCKAGSGGPLICD-----GVLQGITSWGTPCGEPMPG 241

```



QY 253 VYTRVSHFPLWIR 265  
Db 242 VYTKLNKFTSWIK 254			
RESULT 47  
B26823  
pancreatic elastase II (EC 3.4.21.7) A precursor - human  
C/Species: Homo sapiens (man)  
C/Date: 16-Aug-1988 #sequence revision 16-Aug-1988 #text\_change 20-Jun-2000  
C/Accession: B26823, A27432, A41431, S34491  
R/Kawashima, I., Tani, T., Shimoda, K., Takiguchi, Y.  
DNA 6, 163-172, 1987  
A/Title: Characterization of pancreatic elastase II cDNAs: two elastase II mRNAs are expressed  
A/Reference number: A90958; MUID:87217962; PMID:3646943  
A/Accession: B26823  
A/Molecule type: mRNA  
A/Residues: 1-269 <RAW>  
A/Cross-references: GB:M16652; NID:g182057; PIDN:AAA52380.1; PID:g182058  
R/Fletcher, T.S., Shen, W.P., Largman, C.  
Biochemistry 26, 7256-7261, 1987  
A/Title: Primary structure of human pancreatic elastase 2 determined by sequence analysis  
A/Reference number: A27432; MUID:88107669; PMID:3427074  
A/Accession: A27432  
A/Molecule type: mRNA  
A/Residues: 1-263 <LE>  
A/Cross-references: GB:M16631; NID:g182022; PIDN:AAA52374.1; PID:g182023  
R/Shirasu, Y., Yoshida, H., Matsuki, S., Takemura, K., Ikeda, N., Shimada, Y., Ozawa, T., J. Biochem. 102, 1555-1563, 1987  
A/Title: Molecular cloning and expression in Escherichia coli of a cDNA encoding human pancreatic elastase II  
A/Reference number: A41431; MUID:88198076; PMID:2834346  
A/Accession: A41431  
A/Molecule type: mRNA  
A/Residues: 1-201, 'V', 203-269 <SH>  
A/Cross-references: GB:D00236; NID:g219619; PIDN:BAA00165.1; PID:g219620  
A/Note: The authors translated the codon GTG for residue 202 as Cys  
R/Moulard, M., Michon, T., Kerfelec, B., Chapus, C.  
FEBS Lett. 261, 179-183, 1990  
A/Title: Further studies on the human pancreatic binary complexes involving precarboxypeptidase A  
A/Reference number: S08253; MUID:90169111; PMID:2307232  
A/Accession: S34491  
A/Molecule type: protein  
A/Residues: 'X', 18-50 <MOU>  
C/Genetics:  
A/Gene: GDB:ELAI  
A/Cross-references: GDB:119866; OMIM:130120  
A/Map position: 12pter-12qter  
A/Superfamily: trypsin; trypsin homology  
C/Keywords: hydrolase; pancreas; serine proteinase  
F/1-16/Domain: signal sequence #status predicted <SIG>  
F/17-28/Domain: propeptide #status predicted <PRO>  
F/29-262/Product: pancreatic elastase IIA #status predicted <MAT>  
F/29-262/Domain: trypsin homology <TRY>  
F/73,121,216/Active site: His, Asp, Ser #status predicted

Query March 24.78; Score 372; DB 2; Length 269;  
Best Local Similarity 34.0%; Pred. No. 7.9e-45;  
Matches 90; Conservative 46; Mismatches 101; Indels 28; Gaps 10;

QY 13 CGQKTLPRF-KITGGFTTETIQPFAAIRRHRRGGSVTVYCGSLSPCWISATGCF 71  
Db 17 CGDFTYPPYTVRVGGEARPNWPQVSL-QYSSNGKWHTTCGSLIANSWLTAAHCI 75			
QY 72 IDYPKKEDYIVLGRSLNLTQGMKFEVENILKDYSDATLAHNDIALKIRSK-- 129			
Db 76 ---SSSTRYRVGLGRHNLVAESGSLAVSVSKIVHKDMSNOISKGNDAALKIANPVS 132			
QY 130 ----EGRCQPSRTIQTICLPMSYNDPQGTSCIEITFGKGNSTDYLYPEQLKMTWVKL 184			
Db 133 LTDKIQLACLPAGTI-----LPNNY-----PCYVTGWGR-LQTNGAVDVLQGRLLV 180			
QY 185 ISHRECOOPYHYGSEVTKMLCAADPQWKTDSCQDGGGLVC-SLQGRMTLTGIVSWGR 243

Db 181 VDYATCSSAWGSSVKTSMICAGG-DGVISSCNGDSGGLNCOASDGRWQVHGIVSFGS 239  
QY 244 --GCALKDKPGVYTVRSHFPLWIR 265			
Db 240 RLGCNYHYKPSVFTVRSVYIDWINS 264			
RESULT 48  
PLBO  
Plasmin (EC 3.4.21.7) precursor - bovine  
C/Species: Bos primigenius taurus (cattle)  
C/Accession: S45046; A25835; A45961; S03735  
R/Berglund, L., Andersen, M.D., Petersen, T.E.  
submitted to the EMBL Data Library, May 1994  
A/Description: Cloning and characterization of the bovine plasminogen cDNA.  
A/Reference number: S45046  
A/Accession: S45046  
A/Molecule type: mRNA  
A/Residues: 1-812 <BER>  
A/Cross-references: EMBL:X79402; NID:g494962; PIDN:CAA55939.1; PID:g494963  
A/Note: it is uncertain whether Met-1 or Met-8 is the initiator  
R/Schaller, J., Moser, P.W., Dannegger-Muller, G.A.K., Rosselet, S.J., Kampfer, J., Ri  
Eur. J. Biochem. 149, 267-278, 1985  
A/Title: Complete amino acid sequence of bovine plasminogen. Comparison with h  
A/Reference number: A25835; MUID:85203906; PMID:3846532  
A/Accession: A25835  
A/Molecule type: protein  
A/Residues: 27-334, 'D', 336-515, 'H', 517-554, 'L', 556-812 <SCH>  
R/Malinowski, D.P., Sadler, J.E., Davie, E.W.  
Biochemistry 23, 4243-4250, 1984  
A/Title: Characterization of a complementary deoxyribonucleic acid coding for  
A/Reference number: I45961; MUID:85023311; PMID:6148961  
A/Accession: I45961  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 706-743, 'R', 745-812 <MAL>  
A/Cross-references: GB:K02935; NID:g163551; PIDN:AAA30714.1; PID:g163552  
R/Brunisholz, R.A., Lerch, P.G., Schaller, J., Rickli, E.E., Legier, W., Mann  
Eur. J. Biochem. 114, 465-470, 1981  
A/Title: Comparison of the primary structure of the N-terminal CNBr fragments  
A/Reference number: S03735; MUID:81212097; PMID:7238497  
A/Accession: S03735  
A/Molecule type: protein  
A/Residues: 27-83 <BRU>  
C/Function:  
A/Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a  
A/Pathway: fibrinolysis  
C/Superfamily: plasmin; kringle homology; plasminogen-related protein precursor  
C/Keywords: duplication; fibrinolysis; glycoprotein; hydrolase; kidney; kringle  
F/1-26/Domain: signal sequence #status predicted <SIG>  
F/8-103/Domain: plasminogen-related protein precursor homology <PLPH>  
F/27-812/Product: plasminogen #status experimental <PRO>  
F/27-103/Domain: activation peptide #status experimental <APT>  
F/104-583,584-812/Product: plasmin #status experimental <MAT>  
F/104-583/Domain: plasmin chain A #status experimental <ACH>  
F/110-188/Domain: kringle homology <KR1>  
F/192-269/Domain: kringle homology <KR2>  
F/282-359/Domain: kringle homology <KR3>  
F/384-461/Domain: kringle homology <KR4>  
F/485-564/Domain: kringle homology <KR5>  
F/584-812/Domain: plasmin chain B #status experimental <BCH>  
F/584-805/Domain: trypsin homology <TRY>  
F/556-80,60-68,110-188,131-171,159-183,192-269,195-323,213-252,241-264,282-359,  
bonds: #status predicted  
F/315/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F/365/Binding site: carbohydrate (Ser) (covalent) #status experimental  
F/624,667,762/Active site: His, Asp, Ser #status predicted



Query Match 24.7%; Score 372; DB 1; Length 812;  
 Best Local Similarity 35.2%; Pred. No. 2.9e-24;  
 Matches 93; Conservative 34; Mismatches 113; Indels 24; Gaps 6;  
 QY 5 PPEELKFOGQKTLPR---FKIIGCEFTTIENQPFPAIYRRHGGSVTVCGSLISP 61  
 DB 562 PQCESFPQGPKEPKVCKSGRIVGCVSKPHSPWQVSLRSSR-----HFCGGTLIS 616  
 QY 62 CWVISATHCFIDYPPKKEDYIVVLGRSLNSNTQGMKEVENILHKDYSADTLAHNDI 121  
 DB 617 KWLTAACHLNDILALSFKVLGAHNEKVRQEQVEIPVSELFRPQQA-----DI 668  
 QY 122 ALLKIRSEKGRCAQPSRTIOTICLPSMNDPQGTSCBITGFKENSTDYLYPE-QLKMT 180  
 DB 669 ALLKL-----SRPAITIKVEIPACLPFPNMYAARTECVITGWGETQTG---FGEGLLEKA 721  
 QY 181 VVKLISHRECOQPHYGVSEVTKMLCAADPQWKTDSCQDGGGLVCSLQGRMTLITGVS 240  
 DB 722 HLPVLENKVRNEVLDGRVKTELCAHGLIGTDSQDGGGLVCFCKKXILQGVTS 781  
 QY 241 WGRGALKDKPGVYTRVSHFLPMI 264  
 DB 782 WGLGARPKNKPGVYVRVSPYVPMI 805

## RESULT 49

## KYSOA

chymotrypsin (EC 3.4.21.1) A precursor - bovine  
 N;Alternate names: chymotrypsinogen A  
 C;Species: Bos primigenius taurus (cattle)  
 C;Date: 07-May-1981 #sequence revision 07-May-1981 #text\_change 07-May-1999  
 C;Accession: A90235; A93158; S29650; A00952  
 R;Brown, J.R.; Hartley, B.S.  
 Biochem. J. 101, 214-228, 1966  
 A;Title: Location of disulphide bridges by diagonal paper electrophoresis. The disulphid  
 A;Reference number: A90235; MUID:67181721; PMID:5971783  
 A;Accession: A90235  
 A;Molecule type: protein  
 A;Residues: 1-101; N', 103-245 <BRO>  
 R;Blow, D.M.; Birktoft, J.J.; Hartley, B.S.  
 Nature 221, 337-340, 1969  
 A;Title: Role of a buried acid group in the mechanism of action of chymotrypsin.  
 A;Reference number: A93158; MUID:69106266; PMID:5764436  
 A;Contents: annotation; revision to residue 102  
 R;Neloun, B.; Klueh, I.; Kostka, V.; Moravsek, L.; Prusik, Z.; Vanacek, J.; Keil, B.; Sorn  
 Biochim. Biophys. Acta 130, 543-546, 1966  
 A;Title: Covalent structure of bovine chymotrypsinogen A.  
 A;Reference number: A90572; MUID:67183948; PMID:5972866  
 A;Accession: A93158  
 A;Molecule type: protein  
 A;Residues: 1-101; N', 103-245 <MEL>  
 R;Cutruzzola, F.; Ascenzi, P.; Barra, D.; Bolognesi, M.; Menegatti, E.; Sarti, P.; Schne  
 Biochim. Biophys. Acta 1161, 201-208, 1993  
 A;Title: Selective oxidation of Met-192 in bovine alpha-chymotrypsin. Effect on catalyti  
 A;Reference number: S29650; MUID:93160238; PMID:8431470  
 A;Accession: S29650  
 A;Molecule type: protein  
 A;Residues: 1-12; 16-27; 149-160; 181-200 <CUT>  
 R;Snillie, L.B.; Hartley, B.S.  
 Biochem. J. 101, 232-241, 1966  
 A;Title: Histidine sequences in the active centres of some 'serine' proteinases.  
 A;Reference number: A90236; MUID:67181723; PMID:5971785  
 A;Contents: annotation; active site  
 R;Birktoft, J.J.; Blow, D.M.; Henderson, R.; Steitz, T.A.  
 Philos. Trans. R. Soc. Lond. B257, 67-76, 1970  
 A;Title: The structure of alpha-chymotrypsin.  
 A;Reference number: A93754  
 A;Contents: annotation; x-ray crystallography  
 C;Comment: Chymotrypsinogens are synthesized in the acinar cells of pancreas.  
 C;Comment: Tryptic cleavage after Arg-15 results in a fully active enzyme (pi-chymotryps  
 della-chymotrypsin; further chymotryptic cleavage liberates the dipeptide Thr-147 and As  
 d Asn-148 directly from chymotrypsinogen, which leads to the degraded form neochymotryps

C;Superfamily: trypsin; trypsin homology  
 C;Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen  
 F;1-245/Product: chymotrypsinogen #status experimental <ZYM>  
 F;1-13, 16-145, 149-245/Product: alpha-chymotrypsin #status experimental <MPT>  
 F;15-238/Domain: trypsin homology <TRY>  
 F;1-122, 42-58, 136-201, 168-182, 191-220/Disulfide bonds: #status experimental  
 F;57, 102, 195/Active site: His, Asp, Ser #status experimental

Query Match 24.6%; Score 371; DB 1; Length 245;  
 Best Local Similarity 32.3%; Pred. No. 8.7e-25;  
 Matches 83; Conservative 52; Mismatches 100; Indels 22; Gaps 7

QY 13 CGQKTLRPRP----KIIGCEFTTIENQPFPAIYRRHGGSVTVYCGSLISPCWVISAT 68  
 DB 1 CGVPAIQPVLSGLSRVINGEEAVPGSPWQVSLQDK-----TGPHFCGGLIINWVVTAA 56  
 QY 69 HCFIDYPPKKEDYIVVLGRSLNSNTQGMKEFEVENILHKDYSADTLAHNDIALKIRS 128  
 DB 57 HCGV-----TSDVVVAGEFPDQSSSEKIQKLAKIYFKNSKYNLSLFI--NNDITLLKLT 110  
 QY 129 KEGRCAQPSRTIOTICLPSMNDPQGTSCBITGFKENSTDYLYPEQLKMTVVKLISHR 188  
 DB 111 ---AASFQTVSAVCLPSASDDFAAGTTCTVTGWLTRYTNANTPDRLQOASLPILSNT 166  
 QY 189 ECOQPHYGVSEVTKMLCAADPQWKTDSCQDGGGLVCSLQGRMTLITGIVSWGRCALK 248  
 DB 167 NCKK--YWGTKIKDAMICAGAS--GVSSCNGDSGGLVCKKNGAWTLVGLVSWGSSCTST 222  
 QY 249 DXPGVYTRVSHFLPMIR 265  
 DB 223 STPGVVARVTALVNVWQ 239

## RESULT 50

## KPHUI

coagulation factor Xia (EC 3.4.21.27) precursor [validated] - human  
 N;Alternate names: antihemophilic factor C; plasma thromboplastin antecedent  
 C;Species: Homo sapiens (man)  
 C;Date: 13-Aug-1986 #sequence revision 26-May-1994 #text\_change 08-Dec-2000  
 C;Accession: A27431; A00920; A37940  
 R;Asakai, R.; Davies, E.W.; Chung, D.W.  
 Biochemistry 26, 7221-7228, 1987  
 A;Title: Organization of the gene for human factor XI.  
 A;Reference number: A27431; MUID:88107663; PMID:2827746  
 A;Accession: A27431  
 A;Molecule type: DNA  
 A;Residues: 1-625 <ASA>  
 A;Cross-references: GB:XL8295  
 A;Note: The sequence shown follows the authors' translation  
 R;Fujikawa, K.; Chung, D.W.; Hendrickson, L.E.; Davie, E.W.  
 Biochemistry 25, 2417-2424, 1986  
 A;Title: Amino acid sequence of human factor XI, a blood coagulation factor wi  
 A;Reference number: A00920; MUID:86243360; PMID:3636155  
 A;Accession: A00920  
 A;Molecule type: mRNA  
 A;Residues: 1-625 <FUU>  
 A;Cross-references: GB:M13142; NID:g182832; PIDN:AAA52487.1; PID:g182833  
 R;McMullen, B.A.; Fujikawa, K.; Davie, E.W.  
 Biochemistry 30, 2056-2060, 1991  
 A;Title: Location of the disulfide bonds in human coagulation factor XI: the p  
 A;Reference number: A37940; MUID:91152017; PMID:1998667  
 A;Accession: A37940  
 A;Molecule type: protein  
 A;Residues: 28-33; 35-49; X', 51-55; X', 57-63; 70-75; X', 77-79; 107-109; X', 111-11  
 ;280-282; X', 284; 285-297; 313-316; X', 318-319; 320-326; X', 328-330; X', 347-349; 3  
 C;Comment: The proenzyme consists of two identical chains linked by one or mor  
 he active site, and a heavy chain, which associates with high molecular weight  
 C;Genetics:  
 A;Gene: GDB:F11  
 A;Cross-references: GDB:119891; OMIM:264900  
 A;Map position: 4q35-q35  
 A;Introns: 19/1; 73/2; 109/1; 162/2; 199/1; 252/2; 289/1; 343/2; 379/1; 435/2;  
 C;Function:

32-13  
 'X', 3  
 isul  
 (MW) X

A>Description: catalyzes the proteolytic activation of coagulation factor IX  
A.Pathway: blood coagulation intrinsic pathway  
C.Superfamily: coagulation factor XI; trypsin homology  
C.Keywords: blood coagulation; duplication; glycoprotein; hemophilia C; homodimer; hydro  
F:1-15/Domain: signal sequence #status predicted <SIG>  
F:19-387/Product: coagulation factor Xla heavy chain #status experimental <HCH>  
F:19-108/Domain: apple repeat <AP1>  
F:109-198/Domain: apple repeat <AP2>  
F:199-288/Domain: apple repeat <AP3>  
F:290-379/Domain: apple repeat <AP4>  
F:388-625/Product: coagulation factor Xla light chain #status experimental <LCH>  
F:388-618/Domain: trypsin homology <TRY>  
F:20-103,514-581,571-599/Disulfide bonds: #status predicted  
F:29/Disulfide bonds: interchain #status experimental  
F:46-76,50-56,110-193,136-165,140-146,200-283,226-255,230-236,291-374,317-346,321-327,38  
F:90,126,353,450/Binding site: carboxydrate (Asn) (covalent) #status predicted  
F:339/Disulfide bonds: interchain #status predicted  
F:387-388/Cleavage site: Arg-Ile (coagulation factor Xla) #status experimental  
F:431,480,575/Active site: His, Asp, Ser #status predicted  
F:491/Binding site: carboxydrate (Asn) (covalent) #status experimental

Query Match 24.6%; Score 370.5; DB 1; Length 625;  
Best Local Similarity 33.5%; Pred. No. 2.9e-24;  
Matches 89; Conservative 49; Mismatches 99; Indels 29; Gaps 11;

QY	12	QCQKTLRPRFKLIIGSEFTTIENQWFAIY-----RRHGGSVTVVCGSLISPCWVIS	56
Db	379	ECTK-IKPR--IVGTASVRGEWPQVTLHTSTPTQRH-----LCGSLIIGNQWILT	428
QY	67	ATHCF--IDYPK-KEDYIVYLGSRSLNNTQGMKFEVENLILHKDYSADTLAHNDIAL	123
Db	429	AAHCFYGVSPKILRVYSGILNQSEIKEDTS---FFGVQEIHHQYKMAESGY--DIAL	483
QY	124	LKIRKEGRCAQPSRIQICLPSPMYNDPQFGTSCEITGFKENSTDIYLPQLKMTVVK	183
Db	484	LKLET----TVNYTDSQRP-CLPSKGRNVIYTDGVTGMYRKLKDKI-QNTLQKAKIP	538
QY	184	LISHRECQOPHYVYGEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVSWGR	243
Db	539	LVINECQK-RYRGKHITHMICAGYREGGKACKGDSGGLSCKHNEVHVLGITSNGE	597
QY	244	GCALKDKPGVYTVVSHFLPWIRSHTK	269
Db	598	GCAQRERPGVYTVVVEVDWILEKTQ	623

Search completed: May 25, 2004, 14:58:31  
Job time : 17.9649 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2004, 14:43:40 ; Search time 10.6433 Seconds  
(without alignments)  
1350.274 Million cell updates/sec

Title: US-09-880-503-5  
Perfect score: 1508  
Sequence: 1 KPSPPELXFPQCGQKTRP.....VSHFLPWIRSHTEKENGAL 276

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1505	99.8	431	1 UROK_HUMAN	P00749 homo sapien
2	1434.5	95.1	433	1 UROK_PAPCY	P16227 papio cynoc
3	1255	83.2	442	1 UROK_PIG	P04185 sus scrofa
4	1188	78.8	433	1 UROK_BOVIN	Q05589 bos taurus
5	1119.5	74.2	432	1 UROK_RAT	P39598 rattus norv
6	1112.5	73.8	433	1 UROK_MOUSE	P06869 mus musculu
7	688.5	45.7	434	1 UROK_CHICK	P5120 gallus gall
8	583	38.7	562	1 TPA_HUMAN	P00750 homo sapien
9	570.5	37.8	559	1 TPA_RAT	P19637 rattus norv
10	565	37.5	566	1 TPA_BOVIN	Q28198 bos taurus
11	564.5	37.4	559	1 TPA_MOUSE	P11214 mus musculu
12	547	36.3	431	1 URTB_DESRO	P98121 desmodus ro
13	546	36.2	394	1 URTG_DESRO	P49150 desmodus ro
14	543	36.0	477	1 URTI_DESRO	P56338 desmodus ro
15	497.5	33.0	653	1 HGFA_MOUSE	P98119 desmodus ro
16	495	32.8	615	1 FA12_HUMAN	P00748 homo sapien
17	493.5	32.7	655	1 HGFA_HUMAN	Q04756 homo sapien
18	485.5	32.2	603	1 FA12_CAVPO	Q04962 cavia porce
19	453	30.0	593	1 FA12_BOVIN	P98140 bos taurus
20	442.5	29.3	875	1 NEIR_HUMAN	P56730 homo sapien
21	425.5	28.2	761	1 NEIR_MOUSE	Q08762 mus musculu
22	423.5	28.1	418	1 HATT_HUMAN	O60235 homo sapien
23	408.5	27.1	343	1 PS88_HUMAN	P16651 homo sapien
24	407	27.0	638	1 KAL_MOUSE	P26262 mus musculu
25	406.5	27.0	811	1 TMS6_MOUSE	Q9dbi0 mus musculu
26	405	26.9	277	1 KUD_HUMAN	Q9ukr3 homo sapien
27	405	26.8	277	1 KUD_MOUSE	Q8iu80 homo sapien
28	404.5	26.7	811	1 TMS6_HUMAN	Q35453 mus musculu
29	402.5	26.7	436	1 HEFS_MOUSE	P03952 homo sapien
30	402	26.7	638	1 KAL_HUMAN	Q9er04 mus musculu
31	398.5	26.4	455	1 TMS5_MOUSE	P05981 homo sapien
32	396.5	26.3	417	1 HEFS_HUMAN	Q90629 gallus gall
33	395.5	26.2	248	1 TRY3_CHICK	

34	394.5	26.2	243	1	TRY1_BOVIN	P00760 bos taurus
35	392.5	26.0	343	1	PLMN_SHEEP	P81286 ovis aries
36	390.5	25.9	247	1	TRY2_BOVIN	Q28463 bos taurus
37	390	25.9	263	1	CTR2_CANFA	P04813 canis faml
38	389.5	25.8	244	1	KLK6_HUMAN	Q92876 homo sapien
39	388.5	25.8	453	1	TMS3_MOUSE	Q8kit0 mus musculu
40	388	25.7	457	1	TMS5_HUMAN	Q9h383 homo sapien
41	386.5	25.6	342	1	PS88_RAT	Q9es87 rattus norv
42	386.5	25.6	416	1	HEPS_RAT	Q05511 rattus norv
43	385	25.5	638	1	KAL_RAT	P14272 rattus norv
44	384.5	25.5	311	1	TRYG_MOUSE	Q9qul7 mus musculu
45	383.5	25.4	790	1	PLMN_PIG	P06867 sus scrofa
46	382	25.3	263	1	CTRB_RAT	P07338 rattus norv
47	380.5	25.2	342	1	PS88_MOUSE	Q9eed1 mus musculu
48	380	25.2	269	1	EL2_PIG	P08419 sus scrofa
49	379.5	25.2	304	1	TRY3_HUMAN	P35030 homo sapien
50	378.5	25.1	264	1	CTEL_HUMAN	P40313 homo sapien
51	377	25.0	263	1	CTRB_HUMAN	P17538 homo sapien
52	377	25.0	812	1	PLMN_MOUSE	P20918 mus musculu
53	376.5	25.0	247	1	TRY1_HUMAN	P07477 homo sapien
54	376	24.9	321	1	TRYG_HUMAN	Q9nrr2 homo sapien
55	375	24.9	250	1	KLKB_HUMAN	Q9ubx7 homo sapien
56	374.5	24.8	261	1	KLK6_MOUSE	P15947 mus musculu
57	374	24.8	271	1	EL2_RAT	P00774 rattus norv
58	373.5	24.8	247	1	TRY2_CANFA	P08872 canis faml
59	373.5	24.8	248	1	TRY1_CHICK	Q90627 gallus gall
60	372.5	24.7	244	1	TRY2_XENLA	P70059 xenopus lae
61	372.5	24.7	454	1	TMS3_HUMAN	P57727 homo sapien
62	372	24.7	261	1	KLK3_MOUSE	P00756 mus musculu
63	372	24.7	269	1	EL2A_HUMAN	P08217 homo sapien
64	372	24.7	275	1	TRYT_PIG	Q9n2d1 sus scrofa
65	372	24.7	422	1	DESL_HUMAN	Q9ul52 homo sapien
66	372	24.7	812	1	PLMN_BOVIN	P08868 bos taurus
67	371	24.6	245	1	CTRA_BOVIN	P00766 bos taurus
68	371	24.6	269	1	EL2_BOVIN	Q29461 bos taurus
69	370.5	24.6	251	1	KLK6_HUMAN	Q9p0g3 homo sapien
70	370.5	24.6	625	1	FA11_HUMAN	P03951 homo sapien
71	370.5	24.6	810	1	PLAN_HUMAN	P00747 homo sapien
72	370	24.5	245	1	CTRB_BOVIN	P00767 bos taurus
73	370	24.5	855	1	STI4_MOUSE	P56677 mus musculu
74	369.5	24.5	246	1	TRY1_RAT	P00762 rattus norv
75	369.5	24.5	248	1	TRY2_CHICK	Q90628 gallus gall

ALIGNMENTS

RESULT 1	UROK_HUMAN	STANDARD;	PRT;	431 AA.
ID	P00749; Q15844; Q16618; Q969W6;			
AC	21-JUL-1986 (Rel. 01, Created)			
DT	20-MAR-1987 (Rel. 04, Last sequence update)			
DE	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)			
DE	(U-plasminogen activator)			
GN	PLAU.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
NC	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=85215647; PubMed=2987867;			
RA	Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Blasi F.;			
RT	"The human urokinase-plasminogen activator gene and its promoter.";			
RL	Nucleic Acids Res. 13:2759-2771(1985).			
[2]				
RN	SEQUENCE FROM N.A.			
RP	Holmes W.E., Pennica D., Blaber M., Rey M.W., Guenzler W.A.,			
RA	Steffens G.J., Heyneker H.L.;			
RT	"Cloning and expression of the gene for pro-urokinase in Escherichia coli.";			



of two chains, A and B. The high molecular mass form contains a long chain A. Cleavage occurs after residue 155 in the low molecular mass form to yield a short A1 chain.

PHARMACEUTICAL: Available under the name Abbokinase (Abbott). Used in Pulmonary Embolism (PE) to initiate fibrinolysis.

SIMILARITY: Belongs to peptidase family S1.

SIMILARITY: Contains 1 EGF-like domain.

SIMILARITY: Contains 1 kringle domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; X02419; CAA26268.1; -  
EMBL; M15476; AAB61253.1; -  
EMBL; D00244; BAA00175.1; -  
EMBL; D11143; BAA01919.1; -  
EMBL; X02760; CAA26535.1; -  
EMBL; AF377330; AAK53822.1; -  
EMBL; BC013575; AAH13575.1; -  
EMBL; X02226; AAC97138.1; -  
EMBL; X02286; AAB61252.1; -  
EMBL; A21571; CAA01559.1; -  
EMBL; A19397; CAA01390.1; -  
PIR; A00931; UKHU.  
PDB; 1KDU; 31-OCT-93.

Query Match 99.8%; Score 1505; DB 1; Length 431;  
Best Local Similarity 99.6%; Pred. No. 9e-129;  
Matches 275; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPEELKFCGQKTLPRKLIIGBFTTIENQPFPAAYRRHGGSVTVVCGSLIS 60  
Db 156 KPSSPEELKFCGQKTLPRKLIIGBFTTIENQPFPAAYRRHGGSVTVVCGSLIS 215

QY 61 PCWVLSATCFIDYPKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSDTLAHND 120  
Db 216 PCWVLSATCFIDYPKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSDTLAHND 275

QY 121 IALLKRSKEGRCAQPSRTIQTICLPSMYNDPQFTSCBITGKENSITDLYLPQLKWT 180  
Db 276 IALLKRSKEGRCAQPSRTIQTICLPSMYNDPQFTSCBITGKENSITDLYLPQLKWT 335

QY 181 VVKLISHRCQPHYVGSVTTMLCAADPQKNTSCQDGGPLVCSLQGSMTLTGIVS 240  
Db 336 VVKLISHRCQPHYVGSVTTMLCAADPQKNTSCQDGGPLVCSLQGSMTLTGIVS 395

QY 241 WGRGCALKDKPGVYTVVSHFLPWIRSHTKENGLAL 276  
Db 396 WGRGCALKDKPGVYTVVSHFLPWIRSHTKENGLAL 431

RESULT 2  
UROK\_PAPCY STANDARD; PRT; 433 AA.  
AC P16227;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DE 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)  
DE (U-plasminogen activator).  
GN PLAU.  
OS Papio cynocephalus (Yellow baboon).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecoidea; Papio.  
OX NCBI\_TaxID=9556;  
RN [1]  
RP SEQUENCE FROM N.A.

TISSUE=Thoracic aorta;  
MEDLINE=90287734; PubMed=2113276;  
AU Y.P.T., Wang T.W., Clowes A.W.;  
RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminogen activator";  
RL Nucleic Acids Res. 18:3411-3411(1990).  
CC -|- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.  
CC -|- SUBUNIT: Found in high and low molecular mass forms. Each consists of two chains, A and B. The high molecular mass form contains a long chain A. Cleavage occurs after residue 155 in the low molecular mass form to yield a short A1 chain (By similarity).  
CC -|- SIMILARITY: Belongs to peptidase family S1.  
CC -|- SIMILARITY: Contains 1 EGF-like domain.  
CC -|- SIMILARITY: Contains 1 kringle domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; X51935; CAA36200.1; -  
PIR; S14687; UKBAY.  
DR HSP; P00749; ILMM.  
DR MEROS; S01231; -  
DR InterPro; IPR006003; Cys\_Ser\_trypsin.  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR006210; EGF.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR008293; Pept\_S1A\_uPA.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00031; kringle; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR Pfam; PIRSF001144; Urk\_plasm\_act; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00181; EGF; 1.  
DR SMART; SM00330; KR; 1.  
DR SMART; SM00020; Tryp\_Spc; 1.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
DR PROSITE; PS50026; EGF\_3; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS00070; KRINGLE\_2; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein; Kringle; EGF-like domain; Zymogen; Signal.  
FT SIGNAL 1 20  
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.  
FT CHAIN 21 176 CHAIN A (BY SIMILARITY).  
FT CHAIN 155 176 SHORT A CHAIN (A1) (BY SIMILARITY).  
FT CHAIN 178 433 CHAIN B (BY SIMILARITY).  
FT DOMAIN 26 62 EGF-LIKE.  
FT DOMAIN 69 150 KRINGLE.  
FT DOMAIN 151 177 CONNECTING PEPTIDE.  
FT DOMAIN 178 433 SERINE PROTEASE.  
FT DISULFID 30 38 BY SIMILARITY.  
FT DISULFID 32 50 BY SIMILARITY.  
FT DISULFID 52 61 BY SIMILARITY.  
FT DISULFID 167 298 INTERCHAIN (BY SIMILARITY).  
FT DISULFID 208 224 BY SIMILARITY.  
FT DISULFID 216 287 BY SIMILARITY.  
FT DISULFID 315 384 BY SIMILARITY.  
FT DISULFID 347 363 BY SIMILARITY.  
FT DISULFID 374 402 BY SIMILARITY.  
FT ACT\_SITE 223 223 CHARGE RELAY SYSTEM.

FT ACT SITE 274 274 CHARGE RELAY SYSTEM.  
FT ACT SITE 378 378 CHARGE RELAY SYSTEM.  
FT CARBOHYD 324 324 N-LINKED (GLCNAC. . .) (BY SIMILARITY).  
SQ SEQUENCE 433 AA; 48595 MW; 816D22DFDCC9792 CRC64;  
  
Query Match 95.1%; Score 1434.5; DB 1; Length 433;  
Best Local Similarity 92.8%; Pred. No. 2.2e-122;  
Matches 259; Conservative 13; Mismatches 4; Indels 3; Gaps 1;  
  
QY 1 KPSSPEELKFOGQKTLPRFKIIGBFTTIENOPWFAAIYRRHGGSVTVVCGSLIS 60  
DB 155 KPSSPEELKFOGQKTLPRFKIIGBFTTIENOPWFAAIYRRHGGSVTVVCGSLIS 214  
QY 61 PCWVSATHCFIDYPKEDYIVYVGRSLNSNTQEMKFEVENLILHKDYSADTLAHND 120  
DB 215 PCWVSATHCFIDYPKEDYIVYVGRSLNSNTQEMKFEVENLILHKDYSADTLAHND 274  
QY 121 IALLKRSKEGRCQAQPSRTIQTICLPSMYNDPQ---FGTSCEITGFGKENSVDLYLPOL 177  
DB 275 IALLKRSKEGRCQAQPSRTIQTICLPSMYNDPQ---FGTSCEITGFGKENSVDLYLPOL 334  
QY 178 KMTVVKLISHRECOQPHYVGEVTTKVLCAADPQWKTDSCQDGGPLVCSLQGRMTLTG 237  
DB 335 KMTVVKLISHRECOQPHYVGEVTTKVLCAADPQWKTDSCQDGGPLVCSLQGRMTLTG 394  
QY 238 IVSWGRCALKDKPGVYTVSHLPLWIRSHTEENGLAL 276  
DB 395 IVSWGRCALKDKPGVYTVSHLPLWIRSHTEENGLAL 433  
  
RESULT 3  
UD UROK PIG STANDARD; PRT; 442 AA.  
AC P04185;  
DT 20-MAR-1987 (Rel. 04, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)  
DE (U-plasminogen activator).  
GN PLAUF.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9923;  
RN [1] \_  
RN [2] \_  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RX MEDLINE=85087954; PubMed=6096832;  
RA Nagamine Y., Pearson D., Altus M.S., Reich E.;  
FT "cDNA and gene nucleotide sequence of porcine plasminogen activator."  
RL Nucleic Acids Res. 12:9525-9541(1984).  
RN [2]  
RP REVISION TO 241.  
RA Nagamine Y.;  
RL Submitted (DEC-1986) to the EMBL data bank.  
CC -|- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
CC plasminogen to form plasmin.  
CC -|- SIMILARITY: Belongs to peptidase family S1.  
CC -|- SIMILARITY: Contains 1 EGF-like domain.  
CC -|- SIMILARITY: Contains 1 kringle domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL; X01548; CAA25806.1; -;  
DR EMBL; X02724; CAA26511.1; -;  
DR PIR; A00932; UKPG.  
DR HSSP; P00749; 1KDU.  
  
MEROPS; S01.231; -;  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR008293; Pept\_S1A\_UPA.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00051; Kringle; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PIRSF; PIRSF001144; Uro\_kinase; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SMC0030; Tryp\_Spc; 1.  
DR PROSITE; PS00020; EGF\_1; 1.  
DR PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
DR PROSITE; PS50026; EGF\_3; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS50070; KRINGLE\_2; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
KW Kringle; EGF-like domain; Zymogen; Signal.  
FT SIGNAL 1 20  
FT CHAIN 21 442 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.  
FT CHAIN 21 188 CHAIN A (BY SIMILARITY).  
FT CHAIN 190 442 CHAIN B (BY SIMILARITY).  
FT DOMAIN 29 65 EGF-LIKE.  
FT DOMAIN 72 153 KRINGLE.  
FT DOMAIN 154 189 CONNECTING PEPTIDE.  
FT DOMAIN 190 442 SERINE PROTEASE.  
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .).  
FT DISULFID 33 41 BY SIMILARITY.  
FT DISULFID 35 53 BY SIMILARITY.  
FT DISULFID 55 64 BY SIMILARITY.  
FT DISULFID 179 310 INTERCHAIN (BY SIMILARITY).  
FT DISULFID 220 236 BY SIMILARITY.  
FT DISULFID 228 299 BY SIMILARITY.  
FT DISULFID 324 393 BY SIMILARITY.  
FT DISULFID 356 372 BY SIMILARITY.  
FT DISULFID 383 411 BY SIMILARITY.  
FT ACT SITE 235 235 CHARGE RELAY SYSTEM.  
FT ACT SITE 286 286 CHARGE RELAY SYSTEM.  
FT ACT SITE 387 387 CHARGE RELAY SYSTEM.  
FT CONFLICT 241 241 Q -> H (IN REF. 1; CAA25806).  
FT CONFLICT 242 242 Q -> H (IN REF. 1; CAA26511).  
FT CONFLICT 288 288 A -> GS (IN REF. 1; CAA25806).  
SQ SEQUENCE 442 AA; 49116 MW; EE32FCEFS01321EE CRC64;  
  
Query Match 83.2%; Score 1255; DB 1; Length 442;  
Best Local Similarity 82.1%; Pred. No. 4.1e-106;  
Matches 225; Conservative 23; Mismatches 26; Indels 0; Gaps 0.  
  
QY 2 PSSPEELKFOGQKTLPRFKIIGBFTTIENOPWFAAIYRRHGGSVTVVCGSLISP 61  
DB 168 PSSPEELKFOGQKTLPRFKIIGBFTTIENOPWFAAIYRRHGGSVTVVCGSLISP 227  
QY 62 CWVSATHCFIDYPKEDYIVYVGRSLNSNTQEMKFEVENLILHKDYSADTLAHNDI 121  
DB 228 CWVSATHCFIDYPKEDYIVYVGRSLNSNTQEMKFEVENLILHKDYSADTLAHNDI 287  
QY 122 ALLKRSKEGRCQAQPSRTIQTICLPSMYNDPQ---FGTSCEITGFGKENSVDLYLPOLXMTV 181  
DB 288 ALLKRSKEGRCQAQPSRTIQTICLPSMYNDPQ---FGTSCEITGFGKENSVDLYLPOLXMTV 347  
QY 182 VKLISHRECOQPHYVGEVTTKVLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVSW 241  
DB 348 VKLISHRECOQPHYVGEVTTKVLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVSW 407  
QY 242 GRGCAKDKPGVYTVSHLPLWIRSHTEENGLA 275

Db 408 GRECAMKDPGVYTRVSRFLTWHTHVGGENGLA 441

# RESULT 4

UOK\_BOVIN STANDARD; PRT; 433 AA.

AC Q05589; Q28209; (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)

DE (U-plasminogen activator).

GN PLAU.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.

OX NCBI\_TaxID=9913;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Kortic endothelium;

RX MEDLINE=93216119; PubMed=8385052;

RA Kraetzschmar J., Haendler B., Kojima S., Rifkin D.B.,

RA Schleuning W.-D.;

RT "Bovine urokinase-type plasminogen activator and its receptor:

RT cloning and induction by retinoic acid.";

RL Gene 125:177-183(1993).

RN [2]

RP SEQUENCE OF 12-433 FROM N.A.

RC TISSUE=Kidney;

RA Ravn P., Berglund L., Petersen T.E.;

RT "Cloning and characterization of the bovine plasminogen activators uPA

RT and tPA.";

RL Int. Dairy J. 5:605-617(1995).

CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-val bond in

CC plasminogen to form plasmin.

CC -1- INDUCTION: By retinoic acid.

CC -1- SIMILARITY: Belongs to peptidase family S1.

CC -1- SIMILARITY: Contains 1 EGF-like domain.

CC -1- SIMILARITY: Contains 1 kringle domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; L03546; AAA51419.1; -;

DR EMBL; X85801; CAA59796.1; -;

DR PIR; JN0560; JN0560.

DR HSSP; P00749; 1LMW.

DR MEROPS; S01.231; -;

DR InterPro; IPR009003; Cys Ser trypsin.

DR InterPro; IPR006209; EGF-like.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR008293; Pept\_S1A\_uPA.

DR InterPro; IPR001254; Peptidase\_S1.

DR InterPro; IPR001314; Peptidase\_S1A.

DR Pfam; PF00051; kringle; 1.

DR Pfam; PF00089; trypsin; 1.

DR PIRSF; PIRSF001144; Urk plasm act; 1.

DR PRINTS; PR00722; CHYMOTRYPSIN.

DR PRINTS; PR00018; KRINGLE.

DR ProDom; PD000395; Kringle; 1.

DR SMART; SM00130; KR; 1.

DR SMART; SM00020; Tryp\_SPC; 1.

DR PROSITE; PS00022; EGF\_1; 1.

DR PROSITE; PS01186; EGF\_2; FALSE\_NEG.

DR PROSITE; PS00026; EGF\_3; 1.

DR PROSITE; PS00021; KRINGLE\_1; 1.

DR PROSITE; PS50070; KRINGLE\_2; 1.

DR PROSITE; PS50240; TRYPSIN\_DOM; 1.

DR PROSITE; PS00134; TRYPSIN\_HIS; 1.

DR PROSITE; PS00135; TRYPSIN\_SER; 1.

KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;

KW Kringle; EGF-like domain; Signal; Zymogen.

FT SIGNAL 1 20 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.

FT CHAIN 21 433 CHAIN A (BY SIMILARITY).

FT CHAIN 21 179 CHAIN B (BY SIMILARITY).

FT CHAIN 181 433 EGF-LIKE.

FT DOMAIN 29 65 KRINGLE.

FT DOMAIN 72 153 CONNECTING PEPTIDE.

FT DOMAIN 154 180 SERINE PROTEASE.

FT DOMAIN 181 433 BY SIMILARITY.

FT DISULFID 33 41 BY SIMILARITY.

FT DISULFID 35 53 BY SIMILARITY.

FT DISULFID 55 64 BY SIMILARITY.

FT DISULFID 170 301 INTERCHAIN (BY SIMILARITY).

FT DISULFID 211 227 BY SIMILARITY.

FT DISULFID 219 290 BY SIMILARITY.

FT DISULFID 315 384 BY SIMILARITY.

FT DISULFID 347 363 BY SIMILARITY.

FT DISULFID 374 402 BY SIMILARITY.

FT ACT\_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT\_SITE 277 277 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT\_SITE 378 378 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT CONFLICT 189 189 A -> T (IN REF. 2).

SQ SEQUENCE 433 AA; 48730 MW; 4DE1B8D4DA47027A CRC64;

Query Match 78.8%; Score 1188; DB 1; Length 433;

Best Local Similarity 76.4%; Pred. No. 4.7e-100;

Matches 210; Conservative 32; Mismatches 33; Indels 0; Gaps 0

QY 2 PSSPPEELKFCGQKTLRPFKIIGGFTTIENQFWFAAIYRRHGGSVTVCGGSLISP 61

DB 159 PSSPREKEEFQCGOKALRPFKIVGGQVTNAENQFWFAAIYRRHGGSVTVCGGSLISP 218

QY 62 CWVLSATHCFIDYPKEDYIVYLGSRSLNSNTQEMKFEVENLILHKYSDATLAHNDI 121

DB 219 CWVLSATHCFIDHPKKNYIVYLGQSLNSDTRGEMQFEVEKLILHEDYSAESLAHNDI 278

QY 122 ALLKIRSEKGRCAOPSRITQICLPISMYNDPQFQTSCEITGFGKENSTLYLPEQLKMTY 181

DB 279 ALLKIRSRGCAOPSSRISQICLPPEHDAHSRCEITGFGKENPFDYRSDLNKTF 338

QY 182 VKLISHRECOHPHYGVEVTYMLCAADPQWKTSCQDGGGLVCSLOGRMILTGIVSW 241

DB 339 VSLVSHVEVCQOPHYGAEVTDKMLCAADPQWETSCQDGGGLVCTIQGRLLTGIVSW 398

QY 242 GRGCALDKPGVYTRVSRFLTWHTHVGGENGLA 276

DB 399 GRDCAMKYKPGVYTRVSRFLTWHTHVGGENGLA 433

## RESULT 5

UOK\_RAT STANDARD; PRT; 432 AA.

AC P29538.

DT 01-APR-1993 (Rel. 25, Created)

DT 01-APR-1993 (Rel. 25, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)

DE (U-plasminogen activator).

GN PLAU.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.

OX NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Fischer 344;

RX MEDLINE=92233409; PubMed=1568219;

RA Henderson B.R., Tansey W.P., Phillips S.M., Ramshaw I.A.,

RA Keifford R.F.;

RT "Transcriptional and posttranscriptional activation of urokinase  
RT plasminogen activator gene expression in metastatic tumor cells."  
RL Cancer Res. 52:2489-2496(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Kidney;  
RA Rabbani S.A.;  
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.  
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
CC plasminogen to form plasmin.  
CC -!- SUBUNIT: Found in high and low molecular mass forms. Each consists  
CC of two chains, A and B. The high molecular mass form contains a  
CC long chain A. Cleavage occurs after residue 156 in the low  
CC molecular mass form to yield a short A1 chain (By similarity).  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
CC -!- SIMILARITY: Contains 1 EGF-like domain.  
CC -!- SIMILARITY: Contains 1 kringle domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X63434; CAA45028.1; -;  
CC EMBL; X65651; CAA46601.1; -;  
CC PIR; S24604; S18932.  
CC HSSP; P00749; 1KDU.  
CC MEROPS; S01.231; -;  
CC InterPro; IPR009003; Cys\_Ser\_trypsin.  
CC InterPro; IPR006209; EGF\_Like.  
CC InterPro; IPR006210; IEGF.  
CC InterPro; IPR000001; Kringle.  
CC InterPro; IPR008293; Pept\_S1A\_UA.  
CC InterPro; IPR001254; Peptidase\_S1.  
CC InterPro; IPR001314; Peptidase\_S1A.  
CC Pfam; PF00051; kringle; 1.  
CC Pfam; PF00089; trypsin; 1.  
CC PIRSF; PIRSF001144; Urk\_plasm\_act; 1.  
CC PRINTS; PR00722; CHYMOTRYPSIN.  
CC PRINTS; PR00018; KRINGLE.  
CC ProDom; PD000395; Kringle; 1.  
CC SMART; SM00181; EGF; 1.  
CC SMART; SM00130; KR; 1.  
CC SMART; SM00020; Tryp\_SPC; 1.  
CC PROSITE; PS00022; EGF\_1; 1.  
CC PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
CC PROSITE; PS50026; EGF\_3; 1.  
CC PROSITE; PS00021; KRINGLE\_1; 1.  
CC PROSITE; PS00070; KRINGLE\_2; 1.  
CC PROSITE; PS00240; TRYPSIN\_DOM; 1.  
CC PROSITE; PS00134; TRYPSIN\_HIS; FALSE\_NEG.  
CC PROSITE; PS00135; TRYPSIN\_SER; 1.  
CC Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
KW Kringle; EGF-like domain; Zymogen; Signal.  
FT SIGNAL 1 19 POTENTIAL  
FT CHAIN 20 432 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.  
FT CHAIN 156 177 CHAIN A (BY SIMILARITY).  
FT CHAIN 179 432 SHORT A CHAIN (A1) (BY SIMILARITY).  
FT CHAIN 27 63 CHAIN B (BY SIMILARITY).  
FT DOMAIN 20 151 EGF-LIKE.  
FT DOMAIN 152 178 KRINGLE.  
FT DOMAIN 179 432 CONNECTING PEPTIDE.  
FT DOMAIN 31 39 SERINE PROTEASE.  
FT DISULFID 33 51 BY SIMILARITY.  
FT DISULFID 53 62 BY SIMILARITY.  
FT DISULFID 168 300 INTERCHAIN (BY SIMILARITY).  
FT DISULFID 210 226 BY SIMILARITY.  
FT DISULFID 218 289 BY SIMILARITY.  
FT DISULFID 314 383 BY SIMILARITY.

FT DISULFID 346 362 BY SIMILARITY.  
FT DISULFID 373 401 BY SIMILARITY.  
FT ACT\_SITE 225 225 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 276 276 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 377 377 CHARGE RELAY SYSTEM.  
FT CONFLICT 16 16 N -> H (IN REF. 2).  
FT CONFLICT 24 24 E -> G (IN REF. 2).  
FT CONFLICT 332 332 D -> N (IN REF. 2).  
SQ SEQUENCE 432 AA; 47957 MW; 4EB1B96C716244C8 CRC64;  
  
Query Match 74.2%; Score 1119.5; DB 1; Length 432;  
Best Local Similarity 71.4%; Pred. No. 7.5e-94;  
Matches 197; Conservative 39; Mismatches 39; Indels 1; Gaps 1  
  
QY 1 KPSSPPEELKFOGQKTLRPFRKIIGGEFTTENQFWFAAIYRRHRGGS-VTVVGGSII 59  
Db 156 KSSSTDQGGFGQGGKALRPFRKIYGGFTVVENQFWFAAIYLNKGGSPSFKCGSII 215  
PLAU  
QY 60 SPQWISATHCFIDYPKEDYIVYGRSLNNTGEMKFEVENILHKDYSADTLAHN 119  
Db 216 SPQWASATHCFVNPQKEEYVYVYLGQSKRNSYNFCEMKFEVEQLIHEDFDETLAFN 275  
QY 120 DIALKIRSKGRCAPSRITOTICLPSPYNDPQFGTSCEITGFGKENSTDLYLPEQLKM 179  
Db 276 DIALKIRISTGCAQPSRTIOTICLPFRFGDAPFGSDCEITGFGQESATDYFYPKOLKM 335  
QY 180 TVVKLISHRECOQPHYHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIV 239  
Db 336 SVVKIISHCECKQPHYHYGSEINYNKMLCAADPEWKTDSCQDGGPLICNDGRPTLSGIV 395  
QY 240 SWRGCGALKDKPGVYTRVSHFLPWIRSHTKENGLA 275  
Db 396 SWGSGCAERKNKPGVYTRVSYFLNWIOSHIGENGLA 431  
  
RESULT 6  
UROK\_MOUSE STANDARD; PRT; 433 AA.  
ID UROK\_MOUSE  
AC P08869;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)  
DE (U-plasminogen activator).  
CN PLAU.  
OS Mus musculus (Mouse).  
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
CX NCBI\_TaxID=10090;  
[1]  
RP SEQUENCE FROM N.A.  
EX MEDLINE=95179474; PubMed=2985383;  
RA Belin D., Vassalli J.-D., Combepine C., Godeau F., Nagamine Y.,  
RA Reich E., Kocher H.F., Duvoisin R.M.;  
RT "Cloning, nucleotide sequencing and expression of cDNAs encoding  
RT mouse urokinase-type plasminogen activator.";  
RL Eur. J. Biochem. 148:225-232(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
EX MEDLINE=98163489; PubMed=2831940;  
RA Degen S.J.F., Heckel J.L., Reich E., Degen J.L.;  
RT "The murine urokinase-type plasminogen activator gene.";  
RL Biochemistry 26:8270-8279(1987).  
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
CC plasminogen to form plasmin.  
CC -!- SUBUNIT: Found in high and low molecular mass forms. Each consists  
CC of two chains, A and B. The high molecular mass form contains a  
CC long chain A. Cleavage occurs after residue 156 in the low  
CC molecular mass form to yield a short A1 chain (By similarity).  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
CC -!- SIMILARITY: Contains 1 EGF-like domain.  
CC -!- SIMILARITY: Contains 1 kringle domain.



CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL; X02389; CAA26231.1; -;  
 CC EMBL; M17922; AAA40539.1; -;  
 CC PIR; A29420; UKMS.  
 CC HSP; P00749; 1KDU.  
 CC MEROPS; S01.231; -;  
 CC MGD; MGI:97611; Plau.  
 CC InterPro; IPR009003; Cys Ser trypsin.  
 CC InterPro; IPR006209; EGF\_like.  
 CC InterPro; IPR006210; EGF.  
 CC InterPro; IPR000001; Kringle.  
 CC InterPro; IPR008293; Pept SIA uPA.  
 CC InterPro; IPR001254; Peptidase S1.  
 CC InterPro; IPR001314; Peptidase\_SIA.  
 CC Pfam; PF00051; kringle; 1.  
 CC Pfam; PF00089; trypsin; 1.  
 CC Pfam; PF00089; trypsin; 1.  
 CC PIRSF; PIRSF001144; Urk plasm act; 1.  
 CC PRINTS; PR00722; CHYMOTRYPSIN.  
 CC PRINTS; PR00018; KRINGLE.  
 CC ProDom; PD000395; Kringle; 1.  
 CC SMART; SM00181; EGF; 1.  
 CC SMART; SM00130; KR; 1.  
 CC SMART; SM00020; Tryp\_Spc; 1.  
 CC PROSITE; PS00022; EGF\_1; 1.  
 CC PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
 CC PROSITE; PS00026; EGF\_3; 1.  
 CC PROSITE; PS00021; KRINGLE\_1; 1.  
 CC PROSITE; PS00070; KRINGLE\_2; 1.  
 CC PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 CC PROSITE; PS00134; TRYPSIN\_HIS; FALSE\_NEG.  
 CC PROSITE; PS00135; TRYPSIN\_SER; 1.  
 CC PLASMINOGEN activation; Hydrolase; Serine protease; Glycoprotein;  
 CC Kringle; EGF-like domain; Zymogen; Signal.  
 CC SIGNAL 1 20  
 CC CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.  
 CC CHAIN 21 178 CHAIN A (BY SIMILARITY).  
 CC CHAIN 157 178 SHORT A CHAIN (A1).  
 CC CHAIN 180 433 CHAIN B (BY SIMILARITY).  
 CC DOMAIN 28 64 EGF-LIKE.  
 CC DOMAIN 71 152 KRINGLE.  
 CC DOMAIN 153 179 CONNECTING PEPTIDE.  
 CC DOMAIN 180 433 SERINE PROTEASE.  
 CC DISULFID 32 40 BY SIMILARITY.  
 CC DISULFID 34 52 BY SIMILARITY.  
 CC DISULFID 54 63 BY SIMILARITY.  
 CC DISULFID 169 301 INTERCHAIN (BY SIMILARITY).  
 CC DISULFID 211 227 BY SIMILARITY.  
 CC DISULFID 219 290 BY SIMILARITY.  
 CC DISULFID 315 384 BY SIMILARITY.  
 CC DISULFID 347 363 BY SIMILARITY.  
 CC DISULFID 374 402 BY SIMILARITY.  
 CC ACT SITE 226 226 CHARGE RELAY SYSTEM.  
 CC ACT SITE 277 277 CHARGE RELAY SYSTEM.  
 CC ACT SITE 378 378 CHARGE RELAY SYSTEM.  
 CC SEQUENCE 433 AA; 48268 MW; A99C35F6250443P9 CRC64;

Query Match 73.8%; Score 1112.5; DB 1; Length 433;  
 Best Local Similarity 71.0%; Pred. NO. 3.2e-93;  
 Matches 196; Conservative 39; Mismatches 40; Indels 1; Gaps 1;

QY 1 KPSSPEELKFCGQKTLRPFKIGGFTTIEQNPFAALYRRHGRGS-VTVYVCGGSLI 59  
 D5 157 KPSSVDGQGGCQCGKALRPFKIVGFTTEVEQNPFAALYQKNKGSPSPFKCGSLI 216  
 QY 60 SPCWVISAHTGFDYPKKEDYIVLGRSLNSNTCGEMKFEVNLILHKVSAATLAHN 119

Db 217 SPCWVASAAHCFIQLPKKENVTVVYLQSKSSYNPGEKFEVQLIILHYVTRDSLAYHN 276  
 QY 120 DIALLKIRSEKRCACQPSRTIQTICLPFSMYNDPQGTSCETITGFKENSTDYLYPEQLKM 179  
 Db 277 DIALLKIRSTGTQCAQPSRSIQTICLPFRFTDPPGSDCEITGRKESSESYLYPKNLKM 336  
 QY 150 TVVKLISHRECOQPHYGVSEVTTVMCLNADPQWTKDSQGSGLVCSLQGRWTLTGIV 239  
 Db 337 SVTKLVSHQEQMPPHYGVSEINYYKLCADPEWTKDSCKGSGGLICNIEGRPTLSGIV 396  
 QY 240 SWGRGKALKDKPGVVTVRVSHPELWIRSHKTEENGLA 275  
 Db 397 SWGRGCAEKNGVVTVRVSHFLDNIQSHIGEEKGLA 432

## RESULT 7

## UROK CHICK

ID UROK CHICK STANDARD; PRT; 434 AA.  
 AC P15120;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)  
 DE (U-plasminogen activator).  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OC NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90110185; PubMed=2295632;  
 RA Leslie N.D., Kessler C.A., Bell S.M., Degen J.L.;  
 RT "The chicken urokinase-type plasminogen activator gene";  
 RL J. Biol. Chem. 265:1339-1344(1990).  
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
 CC plasminogen to form plasmin.  
 CC -!- SIMILARITY: Belongs to peptidase family S1.  
 CC -!- SIMILARITY: Contains 1 EGF-like domain.  
 CC -!- SIMILARITY: Contains 1 kringle domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation.  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; J05187; AAA49131.1; -;  
 CC EMBL; J05188; AAA49130.1; -;  
 CC PIR; A35005; A35005.  
 CC HSP; P00763; IDPO.  
 CC MEROPS; S01.231; -;  
 CC InterPro; IPR009003; Cys Ser trypsin.  
 CC InterPro; IPR006209; EGF\_like.  
 CC InterPro; IPR006210; EGF.  
 CC InterPro; IPR000001; Kringle.  
 CC InterPro; IPR008293; Pept SIA uPA.  
 CC InterPro; IPR001254; Peptidase S1.  
 CC InterPro; IPR001314; Peptidase\_SIA.  
 CC Pfam; PF00051; kringle; 1.  
 CC Pfam; PF00089; trypsin; 1.  
 CC Pfam; PF00089; trypsin; 1.  
 CC PIRSF; PIRSF001144; Urk plasm act; 1.  
 CC PRINTS; PR00722; CHYMOTRYPSIN.  
 CC PRINTS; PR00018; KRINGLE.  
 CC ProDom; PD000395; Kringle; 1.  
 CC SMART; SM00181; EGF; 1.  
 CC SMART; SM00130; KR; 1.  
 CC SMART; SM00020; Tryp\_Spc; 1.  
 CC PROSITE; PS00022; EGF\_1; 1.  
 CC PROSITE; PS01186; EGF\_2; 1.  
 CC PROSITE; PS00026; EGF\_3; 1.

DR PROSITE, PS00021; KRINGLE\_1; 1.  
 DR PROSITE, PS00070; KRINGLE\_2; 1.  
 DR PROSITE, PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE, PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE, PS00135; TRYPSIN\_SER; 1.  
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
 KRingle; EGF-like domain; Signal; Zymogen.  
 FT SIGNAL 1 20  
 FT CHAIN 21 434 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.  
 FT CHAIN 21 171 CHAIN A (BY SIMILARITY).  
 FT CHAIN 173 434 CHAIN B (BY SIMILARITY).  
 FT DOMAIN 36 72 EGF-LIKE  
 FT DOMAIN 79 158 KRINGLE  
 FT DOMAIN 159 172 CONNECTING PEPTIDE.  
 FT DOMAIN 173 434 SERINE PROTEASE.  
 FT DISULFID 40 48 BY SIMILARITY.  
 FT DISULFID 42 60 BY SIMILARITY.  
 FT DISULFID 62 71 BY SIMILARITY.  
 FT DISULFID 162 296 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 202 218 BY SIMILARITY.  
 FT DISULFID 210 285 BY SIMILARITY.  
 FT DISULFID 310 379 BY SIMILARITY.  
 FT DISULFID 342 358 BY SIMILARITY.  
 FT DISULFID 369 397 BY SIMILARITY.  
 FT ACT\_SITE 217 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 272 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 373 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT CARBOHYD 228 N-LINKED (GLCNAC ) (POTENTIAL).  
 SQ SEQUENCE 434 AA; 49400 MW; 8881048DD66A55 CRC64;  
 Query Match 45.7%; Score 688.5; DB 1; Length 434;  
 Best Local Similarity 47.4%; Pred. No. 8.2e-55;  
 Matches 128; Conservative 48; Mismatches 89; Indels 5; Gaps 2;  
 QY 2 PSPPBELFQGGKTLRFRKLIIGGFTTIENQFPAALYRHRGSGVYVCGSLSP 61  
 DB 151 PSTIEKERTCGRSFYSFKVIGSGQAEVTPWIAIGFONIM-GTDQLCGSLIDP 209  
 QY 62 CWVISATHCFID----YPKEDYIVYVGRSLNSNTQGMKEFEVENLILHKDYSADTLAH 117  
 DB 210 CWVLTAAHCFYNPTKKQPNKSVKYVFLGSLINTNDEHQVFWDEIISHDPDFTDHTGDN 269  
 QY 118 HNDIALKLRKSGRCAQPSRIQIICLPKSYNDPQFGTSCHITGPKENSTDIYLPQL 177  
 DB 270 DNDIALIRTAGGQCAVESNVTYVCLPEKLNLYDNTWCEIAGYKQNSYDIYAQRL 329  
 QY 178 KMTVWKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTG 237  
 DB 330 MSATVNLISQDDCKNKYDSTRVDNVCAGDPLWETDACKGDSGPGVCEHNGMTLYG 389  
 QY 238 IVSWGRCALKKQKPGVYTVSHFLPWIRSH 267  
 DB 390 IVSWGRCALKKQKPGVYTVRYLWIDSN 419  
 RESULT 8  
 ID TPA\_HUMAN STANDARD; PRT; 562 AA.  
 AC P00750; Q15103;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)  
 DE (t-PA) (t-plasminogen activator) (Alteplase) (Reteplase).  
 GN PLAT.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Melanoma;  
 RX MEDLINE=83115262; PubMed=6337343;

RA Pennica D., Holmes W.E., Kohn W.J., Harkins R.N., Vehar G.A.,  
 RA Ward C.A., Bennett W.F., Yelverton E., Seeburg P.H., Heyneker H.L.,  
 RA Goeddel D.V., Collen D.;  
 RT "Cloning and expression of human tissue-type plasminogen activator  
 RT cDNA in E. coli.";  
 RL Nature 301:214-221(1983).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fetal lung;  
 RX MEDLINE=88262579; PubMed=3133640;  
 RA Sasaki H., Saito Y., Hayashi M., Otsuka K., Niwa M.;  
 RT "Nucleotide sequence of the tissue-type plasminogen activator cDNA  
 RT from human fetal lung cells.";  
 RL Nucleic Acids Res. 16:5695-5695(1988).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88054470; PubMed=2824147;  
 RA Reddy V.B., Garramone A.J., Sasek H., Wei C.-M., Watkins P., Galli J.,  
 RA Hsiung N.;  
 RT "Expression of human uterine tissue-type plasminogen activator in  
 RT mouse cells using BPV vectors.";  
 RL DNA 6:461-472(1987).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86196143; PubMed=3009482;  
 RA Friezner Degen S.J., Rajput B., Reich E.;  
 RT "The human tissue plasminogen activator gene.";  
 RL J. Biol. Chem. 261:6972-6985(1986).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=84298137; PubMed=6089198;  
 RA NY T., Eligh F., Lund B.;  
 RT "The structure of the human tissue-type plasminogen activator gene:  
 RT correlation of intron and exon structures to functional and  
 RT structural domains.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359(1984).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86284200; PubMed=3090401;  
 RA Harris T.J., Patel T., Marston F.A., Little S., Emtage J.S.,  
 RA Oostlander G., Volckaert G., Rombauts W., Billiau A., Somer P.;  
 RT "Cloning of cDNA coding for human tissue-type plasminogen activator  
 RT and its expression in Escherichia coli.";  
 RL Mol. Biol. Med. 3:279-292(1986).  
 RN [7]  
 RP SEQUENCE FROM N.A. (ISOFORM SHORT).  
 RC TISSUE=Umbilical vein;  
 RX MEDLINE=90192129; PubMed=2107528;  
 RA Siebert P.D., Fong K.;  
 RT "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from  
 RT human endothelial cells.";  
 RL Nucleic Acids Res. 18:1086-1086(1990).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.L., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RP "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [9]  
RP SEQUENCE OF 212-361 FROM N.A.  
RX MEDLINE=93159566; PubMed=6572897;  
RA Edlund T., Ny T., Raanby M., Heden L.-O., Palm G., Holmgren E.,  
RA Josephson S.;  
RT "Isolation of cDNA sequences coding for a part of human tissue  
RL plasminogen activator";  
RN Proc. Natl. Acad. Sci. U.S.A. 80:349-352(1983).  
RN [10]  
RP SEQUENCE OF 1-36 FROM N.A.  
RX MEDLINE=85289338; PubMed=3161893;  
RA Fisher R., Waller E.K., Grossi G., Thompson D., Tizard R.,  
RA Schleuning W.-D.;  
RT "Isolation and characterization of the human tissue-type plasminogen  
RL activator structural gene including its 5' flanking region.";  
RN J. Biol. Chem. 260:11223-11230(1985).  
RN [11]  
RP SEQUENCE OF 31-562 FROM N.A.  
RX MEDLINE=91291340; PubMed=1368681;  
RA Itagaki Y., Yasuda H., Morinaga T., Mitsuda S., Higashio K.;  
RT "Purification and characterization of tissue plasminogen activator  
RL secreted by human embryonic lung diploid fibroblasts, IMR-90 cells."  
RN Agric. Biol. Chem. 55:1225-1232(1991).  
RN [12]  
RP SEQUENCE OF 36-562.  
RC TISSUE=Melanoma;  
RX MEDLINE=85000468; PubMed=6433976;  
RA Pohl G., Kaellstroem M., Bergsdorf N., Wallen P., Joernvall H.;  
RT "Tissue plasminogen activator: peptide analyses confirm an indirectly  
RL derived amino acid sequence, identify the active site serine residue,  
RT establish glycosylation sites, and localize variant differences."  
RN Biochemistry 23:3701-3707(1984).  
RN [13]  
RP SEQUENCE OF 33-52 AND 311-330.  
RC TISSUE=Melanoma;  
RX MEDLINE=83209620; PubMed=6682760;  
RA Wallen P., Echl G., Bergsdorf N., Raanby M., Ny T., Joernvall H.;  
RT "Purification and characterization of a melanoma cell plasminogen  
RL activator";  
RN Eur. J. Biochem. 132:681-686(1983).  
RN [14]  
RP STRUCTURE OF CARBOHYDRATES.  
RX MEDLINE=90092112; PubMed=1513186;  
RA Pfeiffer G., Schmidt M., Strube K.-H., Geyer R.;  
RT "Carbohydrate structure of recombinant human uterine tissue  
RL plasminogen activator expressed in mouse epithelial cells."  
RN Eur. J. Biochem. 186:273-286(1989).  
RN [15]  
RP CARBOHYDRATE-LINKAGE SITE THR-96.  
RX MEDLINE=91159408; PubMed=1900431;  
RA Harris R.J., Leonard C.K., Guzzetta A.W., Spellman M.W.;  
RT "Tissue plasminogen activator has an O-linked fucose attached to  
RL threonine-61 in the epidermal growth factor domain."  
RN Biochemistry 30:2311-2314(1991).  
RN [16]  
RP DISULFIDE BONDS IN KRINGLE 2.  
RX MEDLINE=91244765; PubMed=1645336;  
RA Vlahos C.J., Wilhelm O.G., Hassell T., Jaekunas S.R., Bang N.U.;  
RT "Disulfide pairing of the recombinant kringle-2 domain of tissue  
RL plasminogen activator produced in Escherichia coli."  
RN J. Biol. Chem. 266:10070-10072(1991).  
RN [17]  
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF CATALYTIC DOMAIN.  
RX MEDLINE=96200985; PubMed=8613982;  
RA Lamba D., Bauer M., Huber R., Fischer S., Rudolph R., Kohnert U.,  
RA Bode W.;  
RT "The 2.3 A crystal structure of the catalytic domain of recombinant  
RL two-chain human tissue-type plasminogen activator."  
RN J. Mol. Biol. 258:117-135(1996).  
RN [18]

RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF CATALYTIC DOMAIN.  
RX MEDLINE=97449126; PubMed=9305622;  
RA Renatus M., Engh R.A., Stubbs M.T., Huber R., Fischer S., Kohnert U.,  
RA Bode W.;  
RT "Lysine 156 promotes the anomalous proenzyme activity of tPA: X-ray  
RL crystal structure of single-chain human tPA."  
RN EMBO J. 16:4797-4805(1997).  
RN [19]  
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF KRINGLE 2.  
RX MEDLINE=92118803; PubMed=1310033;  
RA de Vos A., Ullsch M.H., Kelley R.F., Padmanabhan K., Tulinsky A.,  
RA Westbrook M.L., Kosiakof A.A.;  
RT "Crystal structure of the kringle 2 domain of tissue plasminogen  
RL activator at 2.4-A resolution."  
RN Biochemistry 31:270-279(1992).  
RN [20]  
RP STRUCTURE BY NMR OF KRINGLE 2.  
RX MEDLINE=90122799; PubMed=2558718;  
RA Byeon I.-J.L., Kelley R.F., Llinas M.;  
RT "1H NMR structural characterization of a recombinant kringle 2 domain  
RL from human tissue-type plasminogen activator."  
RN Biochemistry 28:9350-9360(1989).  
RN [21]  
RP STRUCTURE BY NMR OF KRINGLE 2.  
RX MEDLINE=91200042; PubMed=1901789;  
RA Byeon I.-J.L., Kelley R.F., Llinas M.;  
RT "Kringle-2 domain of the tissue-type plasminogen activator 1H-NMR  
RL assignments and secondary structure."  
RN Eur. J. Biochem. 197:155-165(1991).  
RN [22]  
RP STRUCTURE BY NMR OF KRINGLE 2.  
RX MEDLINE=92106329; PubMed=1762144;  
RA Byeon I.-J.L., Llinas M.;  
RT "Solution structure of the tissue-type plasminogen activator kringle  
RL 2 domain complexed to 6-aminohexanoic acid an antifibrinolytic  
RL drug."  
RN J. Mol. Biol. 222:1035-1051(1991).  
RN [23]  
Query Match 38.7%; Score 583; DB 1; Length 562;  
Best Local Similarity 44.9%; Pred. No. 4.le-45;  
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5  
QY 13 CG-OXKLRPRFKIIGGFTTIENQFWFAAIYRRH-RGGSVTVYCGSLISPCWVISATHC 70  
DB 299 CGLRQYQOPFRINGGLFADIASHPQAAIFAKHRRSPGERFLCGGLISCSWILSAHC 358  
QY 71 FIDYPKEDYIVYLGSRSLNSNTQEMKFEVENILHKDYADTLAHNDTALKIRSK 130  
DB 359 FOERPPPHLTVILGRIVRVVPGSEOKFEVEKIIVHKEFDDDT--YNDIALQLKSDS 416  
QY 131 GRCAQPSRTIOTICLPMSYNDPQFGTSCETIGFGKENSIDLYPEQLKMTVYKLISHREC 190  
DB 417 SRCAGSSVATVCLPPADLQLPDWTCELSGYCKHEALSPFYSERLKEAHVRLVPSRC 476  
QY 191 QOPHYGSEVTVKMLCAAD-----POWKT-DSOCDGSGPLVCSLOGRWTLTGIVSWGRG 244  
DB 477 TSOHLNRTVTDNMLCAGDTRSGGPOANLHDACQDGGPLVCLNDGRMTLVGIISWGLG 536  
QY 245 CALKDKPGVTVRSVHFLPWIRSHTK 269  
DB 537 CGQKDVPGVYTKYNLDWIRDNR 561  
RESULT 9  
TPA RAT  
ID TPA RAT STANDARD; PRT; 559 AA.  
AC P19637;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)  
DE (t-PA) (t-plasminogen activator).

GN PLAT.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89170114; PubMed=3148445;  
RA NY T., Leonardson G., Hueh A.J.W.;  
RT "Cloning and characterization of a cDNA for rat tissue-type  
RT plasminogen activator.";  
RL DNA 7:671-677(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90130448; PubMed=2105315;  
RA Feng P., Ohlsson M., Ny T.;  
RT "The structure of the TATA-less rat tissue-type plasminogen activator  
RT gene. Species-specific sequence divergences in the promoter predict  
RT differences in regulation of gene expression.";  
RL J. Biol. Chem. 265:2022-2027(1990).  
CC -1- FUNCTION: Converts the abundant, but inactive, zymogen plasminogen  
CC to plasmin by hydrolyzing a single Arg-Val bond in plasminogen. By  
CC controlling plasmin-mediated proteolysis, it plays an important  
CC role in tissue remodeling and degradation, in cell migration and  
CC many other physiological events.  
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
CC plasminogen to form plasmin.  
CC -1- SUBUNIT: Heterodimer of chain A and chain B held by a disulfide  
CC bond.  
CC -1- SUBCELLULAR LOCATION: Secreted; extracellular.  
CC -1- PM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER  
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER  
CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.  
CC -1- MISCELLANEOUS: Binds to the kringle structure of the fibrin A  
CC chain. Binding to fibrin enhances its catalytic activity.  
CC -1- SIMILARITY: Belongs to peptidase family S1.  
CC -1- SIMILARITY: Contains 1 EGF-like domain.  
CC -1- SIMILARITY: Contains 1 fibronectin type I domain.  
CC -1- SIMILARITY: Contains 2 kringle domains.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M23697; AAA41812.1; -  
CC EMBL; M31197; AAA42261.1; -  
CC EMBL; M31185; AAA42261.1; JOINED.  
CC EMBL; M31186; AAA42261.1; JOINED.  
CC EMBL; M31187; AAA42261.1; JOINED.  
CC EMBL; M31188; AAA42261.1; JOINED.  
CC EMBL; M31189; AAA42261.1; JOINED.  
CC EMBL; M31190; AAA42261.1; JOINED.  
CC EMBL; M31191; AAA42261.1; JOINED.  
CC EMBL; M31192; AAA42261.1; JOINED.  
CC EMBL; M31193; AAA42261.1; JOINED.  
CC EMBL; M31194; AAA42261.1; JOINED.  
CC EMBL; M31195; AAA42261.1; JOINED.  
CC EMBL; M31196; AAA42261.1; JOINED.  
CC EMBL; A19618; CAA01482.1; -  
CC PIR; A35029; A35029.  
CC HSSP; P00750; 1RTF.  
CC  
CC InterPro; IPR003003; Cys Ser trypsin.  
CC InterPro; IPR006209; EGF-like.  
CC InterPro; IPR000083; Fibinectnl.  
CC InterPro; IPR006210; IEGF.  
CC InterPro; IPR000001; Kringle.  
CC InterPro; IPR001254; Peptidase S1.  
CC InterPro; IPR001314; Peptidase\_S1A.

DR Pfam; PF00008; EGF; 1.  
DR Pfam; PF00039; EGF; 1.  
DR Pfam; PF00031; kringle; 2.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000395; Kringle; 2.  
DR SMART; SM00181; EGF; 1.  
DR SMART; SM00058; EGF; 1.  
DR SMART; SM00130; KR; 2.  
DR SMART; SM00020; Tryp\_Spc; 1.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS01186; EGF\_2; 1.  
DR PROSITE; PS00026; EGF\_3; 1.  
DR PROSITE; PS01253; FIBRONECTIN\_1; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 2.  
DR PROSITE; PS00070; KRINGLE\_2; 2.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
KW Plasma; Kringle; EGF-like domain; Repeat; Signal.  
FT SIGNAL 1 17  
FT PROPEP 18 29  
FT CHAIN 30 559  
FT CHAIN 30 308  
FT CHAIN 309 559  
FT CHAIN 36 78  
FT DOMAIN 79 117  
FT DOMAIN 124 205  
FT DOMAIN 213 294  
FT DOMAIN 309 355  
FT ACT\_SITE 355 355  
FT ACT\_SITE 404 404  
FT ACT\_SITE 510 510  
FT DISULFID 38 68  
FT DISULFID 66 75  
FT DISULFID 83 94  
FT DISULFID 88 105  
FT DISULFID 107 116  
FT DISULFID 124 205  
FT DISULFID 145 187  
FT DISULFID 176 200  
FT DISULFID 213 294  
FT DISULFID 234 276  
FT DISULFID 265 289  
FT DISULFID 297 428  
FT DISULFID 340 356  
FT DISULFID 348 417  
FT DISULFID 442 516  
FT DISULFID 474 490  
FT DISULFID 506 534  
FT CARBOHYD 149 149  
FT CARBOHYD 481 481  
FT CONFLICT 380 380  
SQ SEQUENCE 559 AA; 62903 MW; 7DBD3809C1D1C921 CRC64;  
Query Match 37.8%; Score 570.5; DB 1; Length 559;  
Best Local Similarity 42.4%; Pred. No. 5.5e-44;  
Matches 114; Conservative 47; Mismatches 91; Indels 17; Gaps 6  
QY 13 CQKTLR-PRPKIIGETTTENQWPAAY-RHRGGSVTVYCGGLISPCWISATHC 70  
DB 297 CGLRQYKQPFRIKGLFTDITSHPWQAAPVFNKRSRPFGLCGVLISSCWLSAHC 356  
QY 71 FIDYPKEDYIVYGLGRSLNNTQGMKFEVENILHKDYSADTLAHNDIALKRSKE 130  
DB 357 FVERPPPHLKVVLGRVTVVPGEEQTFEIKIVHKEFDDDT--YNDIALQLRSDS 414  
QY 131 GRCAQPSRTIQICLPMSYNDPQF---GTSCETGFGKENSTDYLPYEQIKMTVVKLI 186

Db 415 SQCAESSVGTACLP-----DPDVQLPDWTECELSGYGKHASSPFFSDRLKEAHRVLYP 470

QY 187 HRECOQPHYGVSEVTKMLCAADP-----QWKIDSQCGSDGGPLVCSLQGRMTLTGVISW 241

Db 471 SSRCTSOHLFNKTIKTNMLCAGTRTGGNQDVHDACQSDGGPLVCMIDKRWMTLLGIISW 530

QY 242 GRGCALKDKPGVYTRVSHPLWIRSHTK 270

Db 531 GLGCGQKQVFGIYTKVNTYLNIDQNMKQ 559

RESULT 10

TPA\_BOVIN STANDARD; PRT; 566 AA.

AC Q28198;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)

GN (t-PA) (t-plasminogen activator).

DE PLAT

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.

OX NCBI\_TaxID=9913;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=kinney;

RA Ravn P., Berglund L., Petersen T.E.;

RT "Cloning and characterization of the bovine plasminogen activators uPA

RT and tPA."

RL Int. Dairy J. 5:605-617(1995).

CC -1- FUNCTION: Converts the abundant, but inactive, zymogen plasminogen to plasmin by hydrolyzing a single Arg-Val bond in plasminogen. By controlling plasmin-mediated proteolysis, it plays an important role in tissue remodeling and degradation, in cell migration and many other physiological events.

CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.

CC -1- SUBUNIT: Heterodimer of chain A and chain B held by a disulfide bond.

CC -1- SUBCELLULAR LOCATION: Secreted; extracellular.

CC -1- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER ARG-314 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.

CC -1- MISCELLANEOUS: Binds to the kringle structure of the fibrin A chain. Binding to fibrin enhances its catalytic activity.

CC -1- SIMILARITY: Belongs to peptidase family S1.

CC -1- SIMILARITY: Contains 1 EGF-like domain.

CC -1- SIMILARITY: Contains 1 fibronectin type I domain.

CC -1- SIMILARITY: Contains 2 kringle domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; X85800; CAA59795.1; --

CC HSP; P00750; 1RTF.

CC MEROPS; S01.232.

CC InterPro; IPR009003; Cys\_Ser\_trypsin.

CC InterPro; IPR006209; EGF\_Like.

CC InterPro; IPR000083; Fibrinctn1.

CC InterPro; IPR006210; IEFG.

CC InterPro; IPR000001; Kringle.

CC InterPro; IPR001254; Peptidase\_S1.

CC InterPro; IPR001314; Peptidase\_S1A.

CC Pfam; PF00008; EGF; 1.

CC Pfam; PF00039; Inl; 1.

DR Pfam; PF00051; kringle; 2.

DR Pfam; PF00089; trypsin; 1.

DR PRINTS; PRO0722; CHYMOTRYPSIN.

DR PRINTS; PRO0018; KRINGLE.

DR ProDom; PD000395; Kringle; 2.

DR SMART; SM00181; EGF; 1.

DR SMART; SM00058; FN1; 1.

DR SMART; SM00130; KR; 2.

DR SMART; SM00020; Tryp\_Spc; 1.

DR PROSITE; PS00022; EGF\_1; 1.

DR PROSITE; PS01186; EGF\_2; 1.

DR PROSITE; PS00026; EGF\_3; 1.

DR PROSITE; PS01253; FIBRONECTIN\_1; 1.

DR PROSITE; PS00021; KRINGLE\_1; 1.

DR PROSITE; PS00070; KRINGLE\_2; 2.

DR PROSITE; PS00240; TRYPSIN\_DOM; 1.

DR PROSITE; PS00134; TRYPSIN\_HIS; 1.

DR PROSITE; PS00135; TRYPSIN\_SER; 1.

KW plasminogen activation; Hydrolase; Serine protease; Glycoprotein; plasma; Kringle; EGF-like domain; Repeat; Signal.

FT SIGNAL 1 21

FT PROPEP 22 33

FT CHAIN 34 566

FT CHAIN 34 314

FT CHAIN 315 566

FT DOMAIN 40 82

FT DOMAIN 83 121

FT DOMAIN 128 209

FT DOMAIN 219 300

FT DOMAIN 315 566

FT ACT\_SITE 361 361

FT ACT\_SITE 410 410

FT ACT\_SITE 517 517

FT DISULFID 42 72

FT DISULFID 70 79

FT DISULFID 87 98

FT DISULFID 92 109

FT DISULFID 111 120

FT DISULFID 128 209

FT DISULFID 149 191

FT DISULFID 180 204

FT DISULFID 219 300

FT DISULFID 240 282

FT DISULFID 271 295

FT DISULFID 303 434

FT DISULFID 346 362

FT DISULFID 354 423

FT DISULFID 448 523

FT DISULFID 480 496

FT DISULFID 513 541

FT CARBOHYD 153 153

FT CARBOHYD 487 487

SQ SEQUENCE 566 AA; 63701 MW; 2EB6BE4E32276C3 CRC64;

Query Match 37.5%; Score 565; DB 1; Length 566;

Best Local Similarity 43.4%; Pred. No. 1.8e-43;

Matches 115; Conservative 41; Mismatches 99; Indels 10; Gaps 4

QY 13 CG-QXLRPRKIIIGGEFTTIENQWFAAIY-RHRGGSVTVYCGSLISPCWVISTHC 70

Db 303 CGLROYKRPQFRKGLFADITSHPQAAIFVKNRSPGERFLCGGILISSCWLSAHC 362

QY 71 FIDYPKEDYIYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHNDLALKIRSK 130

Db 363 PQERYPPHLKVLGRTYRLVPGEEQTFEVEKTIHKFDDDT--YNDIALHLKSDS 420

QY 131 GRCAQPSRTIOTICLPMSYNDPQFCTSCETGFKENSTDYLYPEQLKMTVVKLIHRE 190

Db 421 LTCARESASVETICLPDASLQLPDWTECELSGYGKHSSSPFFSERLKEAHRVLPSSRC 480

QY 191 QPHYGVSEVTKMLCAADPQW-----KTDSQCGSDGGPLVCSLQGRMTLTGVISWGRG 244

481 TQOHFNRTVTNNMLCAGDTSGGSDHTNLHDACQDGGPLVCMKDNHMTLVGLIISWGLG 540

245 CALKDKPGVYTRVSHFLPWRSHTK 269

541 CGRKDPVPGVYTRVNYLDWRDNR 565

RESULT 11

TPA_MOUSE	ID	TPA_MOUSE	STANDARD;	PRT;	559 AA.
AC	PI1214;	Q91VP2;			
DT	01-JUL-1989	(Rel. 11, Created)			
DT	10-OCT-2003	(Rel. 42, Last sequence update)			
DT	15-MAR-2004	(Rel. 43, Last annotation update)			
DE	Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)				
DE	(t-PA) (t-plasminogen activator).				
GN	PLAT.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
NCBI_TaxID=10090;					
RN	[1]				
RP	SEQUENCE FROM N.A.				
RP	MEDLINE=88087303; PubMed=2826484;				
EX	Rickles R.J.; Barrow A.L.; Strickland S.				
RA	"Molecular cloning of complementary DNA to mouse tissue plasminogen				
RT	activator mRNA and its expression during F9 teratocarcinoma cell				
RT	differentiation."				
RL	J. Biol. Chem. 263:1563-1569 (1988).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RP	TISSUE=Mammary gland;				
RC	MEDLINE=22388257; PubMed=12477932;				
EX	Strausberg R.L.; Reinhold E.A.; Grouse L.H.; Derge J.G.				
RA	Klausner R.D.; Collins F.S.; Wagner L.; Shemmen C.M.; Schuler G.D.,				
RA	Altshul S.F.; Zeeberg B.; Buetow K.H.; Schaefer C.F.; Bhat N.K.,				
RA	Hopkins R.F.; Jordan H.; Moore T.; Max S.I.; Wang J.; Hsieh F.,				
RA	Diatchenko L.; Marusina K.; Farmer A.; Rubin G.M.; Hong L.,				
RA	Stapleton M.; Soares M.B.; Bonaldi M.F.; Casavant T.L.; Scheetz T.E.,				
RA	Brownstein M.J.; Usdin T.B.; Toshiyuki S.; Carninci P.; Prange C.,				
RA	Raha S.; Loquellano N.A.; Peters G.J.; Abramson R.D.; Mullahy S.J.,				
RA	Bosak S.A.; McEwan P.J.; McKernan K.J.; Malek J.A.; Gunaratne P.H.,				
RA	Richards S.; Worley K.C.; Hale S.; Garcia A.M.; Gay L.J.; Hulyk S.W.,				
RA	Villaion D.K.; Muzny D.M.; Sodergren E.J.; Lu X.; Gibbs R.A.,				
RA	Fahney J.; Helton E.; Ketterman M.; Madan A.; Rodrigues S.; Sanchez A.,				
RA	Whiting M.; Madan A.; Young A.C.; Shevchenko Y.; Bouffard G.G.,				
RA	Blakesley R.W.; Touchman J.W.; Green E.D.; Dickson M.C.,				
RA	Rodriguez A.C.; Grimwood J.; Schmutz J.; Myers R.M.,				
RA	Butterfield Y.S.N.; Krzyzinski M.I.; Skalska U.; Smallos D.E.,				
RA	Scherer A.; Schein J.F.; Jones S.J.M.; Warra M.A.				
RT	"Generation and initial analysis of more than 15,000 full-length human				
RT	and mouse cDNA sequences."				
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).				
CC	-I- FUNCTION: Converts the abundant, but inactive, zymogen plasminogen				
CC	to plasmin by hydrolyzing a single Arg-Val bond in plasminogen. By				
CC	controlling plasmin-mediated proteolysis, it plays an important				
CC	role in tissue remodeling and degradation, in cell migration and				
CC	many other physiological events				
CC	-I- CATALYTIC ACTIVITY: Specific cleavage of Arg- -Val bond in				
CC	plasminogen to form plasmin.				
CC	-I- SUBUNIT: Heterodimer of chain A and chain B held by a disulfide				
CC	bond.				
CC	-I- SUBCELLULAR LOCATION: Secreted; extracellular.				
CC	-I- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER				
CC	PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER				
CC	ARG-308 CATALYZED BY PLASMIN. TISSUE KALLIKREIN OR FACTOR XA				
CC	-I- MISCELLANEOUS: Binds to the kringle structure of the fibrin A				
CC	chain. Binding to fibrin enhances its catalytic activity.				
CC	-I- SIMILARITY: Belongs to peptidase family S1.				
CC	-I- SIMILARITY: Contains 1 EGF-like domain.				
CC	-I- SIMILARITY: Contains 1 fibronectin type I domain.				
CC	-I- SIMILARITY: Contains 2 kringle domains.				

```
FT DISULFID 442 516 BY SIMILARITY.
FT DISULFID 474 490 BY SIMILARITY.
FT DISULFID 506 534 BY SIMILARITY.
FT CARBOHYD 149 149 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 260 325 G -> A (IN REF. 1).
FT CONFLICT 325 325 P -> A (IN REF. 1).
SQ SEQUENCE 559 AA; 63122 MW; 8CCEB2BDB94514D9 CRC64;

Query Match
Best Local Similarity 37.4%; Score 564.5; DB 1; Length 559;
Matches 114; Conservative 47; Mismatches 91; Indels 17; Gaps 6;

QY 13 CG-QKTLRPRFKIIGSEFTIINQWPAAY-RRHGGSVTVVCGSLISPCWVLSATHC 70
DB 297 CGLRQYRQPRFKIGGLYDITSHFWQAPDFVKNRSPGERFLCGGLVLISSCWVLSAAHC 356
QY 71 FIDYPKKEDYIVVGLGRSLNSNTQGMKEVENLILHKDYSADTLAHNDIALKIRSKS 130
DB 357 FLERPPNHLKVLGRTYRVGEEHCTFEIKYIVHEEFDDET--YNDIALQLRSQS 414
QY 131 GRCAQPSRTIQTCLPSMYNDPQF-----QTSCEITFGKENSTDYLYPEQLKMTVVKLIS 186
DB 415 KQCAQESSVGTACLP-----DNLQLPDWTECELSGYGKHEASSPFFSDELKEAHVRLYP 470
QY 187 HRECQOPHYGVSEVTMLCAADP-----QWKTDSQGDSDGPGVCSLQGRMTLTGIYSW 241
DB 471 SSRCTSQHLNFKVTNNMLCAGDTRSGNQDLHDACQDGGPLVCMINKQMTLTGIISW 530
QY 242 GRGCAKDKPGVYTRVSHLPWIRSHKE 270
DB 531 GLGCGQKDVGVYKVTYNLWDHDMKQ 559

RESULT 12
URTB_DESRO STANDARD; PRT; 431 AA.
AC P98121;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator beta precursor (EC 3.4.21.68) (DSFA
DE beta).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1];
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RN [2];
RN CHARACTERIZATION.
RP MEDLINE=93393059; PubMed=1309059;
RX Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Balduz B., Witt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC -!- FUNCTION: Probably essential to support the feeding habits of this
CC exclusively haematophagous animal. Probable potent thrombolytic
CC agent.
CC -!- CATABOLIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: Monomer.
CC -!- SIMILARITY: Belongs to peptidase family S1.
```

```
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaborative
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M63989; AAA31594.1; --
CC FIR; JS0399; JS0599.
CC HSPF; P98119; IASI.
CC MEROPS; S01.239; --
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR006209; EGF-like.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Peptidase_S1.
CC InterPro; IPR001314; Peptidase_S1A.
CC Pfam; PF00008; EGF; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC PRODOM; PD000395; Kringle; 1.
CC SMART; SM00181; EGF; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00020; Tryp_Spc; 1.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF-2; 1.
CC PROSITE; PS50026; EGF-3; 1.
CC PROSITE; PS50021; KRINGLE_1; 1.
CC PROSITE; PS50070; KRINGLE_2; 1.
CC PROSITE; PS50240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
CC Kringle; EGF-like domain; Signal; Multigene family.
FT SIGNAL 1 36 POTENTIAL.
FT CHAIN 37 431 SALIVARY PLASINOGEN ACTIVATOR BETA.
FT DOMAIN 37 75 EGF-LIKE.
FT DOMAIN 82 163 KRINGLE.
FT DOMAIN 179 431 SERINE PROTEASE.
FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 275 275 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 382 382 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 41 52 BY SIMILARITY.
FT DISULFID 46 63 BY SIMILARITY.
FT DISULFID 65 74 BY SIMILARITY.
FT DISULFID 82 163 BY SIMILARITY.
FT DISULFID 103 145 BY SIMILARITY.
FT DISULFID 134 158 BY SIMILARITY.
FT DISULFID 168 299 BY SIMILARITY.
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 288 BY SIMILARITY.
FT DISULFID 313 388 BY SIMILARITY.
FT DISULFID 345 361 BY SIMILARITY.
FT DISULFID 378 406 BY SIMILARITY.
FT CARBOHYD 139 139 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 431 AA; 48221 MW; 699B5E675B162CBF CRC64;

Query Match 36.3%; Score 547; DB 1; Length 431;
Best Local Similarity 42.9%; Pred. No. 5.4e-42;
Matches 118; Conservative 43; Mismatches 102; Indels 12; Gaps 6

QY 4 SPPELRFQCG-QKTLRPRFKIIGSEFTIINQWPAAYRRHGGSVTVVCGSLISPC 61
DB 159 SVPVCSKATCGLRKYKFPQLHSTGGLFTDITSHFWQAAIFAQNRSSGERFLCGGLISS 218
QY 62 CWVISATHCFID-YPKKEDYIVVGLGRSLNSNTQGMKEVENLILHKDYSADTLAHND 120
```



```
Db 219 CWLTAHCFQERYPPQHLRVV-LGRYVRKPGKEQTFFVEKCIIEHEFDDDT--YNNND 275
Qy 121 IALLKIRSKERCAQPSRTIOTICLPMSYNDPQGTSCETIGKGNSTDYLYPEOLKMT 180
Db 276 IALLQKSGSPQCAQESDSVRAICLPEANLQLPDWTCELSGKGKSSFFSEQLKEG 335
Qy 181 VKLISHRECOQPHYGSEVTTKMLCAADPOWKT-----DSCQDGGGGLVCSLQGRMT 234
Db 336 HVRYPSSRCTSKFLNKTNTNNMLCAGTRSGBEYIPNVHDAQCQDGGGLVCMNDNMT 395
Qy 235 LTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTK 269
Db 396 LGGIISWGVGGEKDFGVYTKVNYLGIWIRDNR 430

RESULT 13
ID URT2 DESRO STANDARD; PRT; 394 AA.
AC P49150;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, last sequence update)
DE 28-FEB-2003 (Rel. 41, last annotation update)
DE Salivary plasminogen activator gamma precursor (EC 3.4.21.68) (DSPA gamma).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kraetschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.B.;
RT "The plasminogen activator family from the salivary gland of the
RL vampire bat Desmodus rotundus: cloning and expression.";
RN Gene 105:229-237(1991).
RP [2]
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RL vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC -!- FUNCTION: Probably essential to support the feeding habits of this
CC exclusively haematophagous animal. Probable potent thrombolytic
CC agent.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: Monomer.
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M63990; AAA31595.1; -
DR PIR; J50600; JS0600.
DR HSSP; P98119; IAS1.
DR MEROPS; S01.239; -.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kringle_S1.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
```

```
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activator; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; Signal; Multigene family.
FT SIGNAL 1 36
FT CHAIN 37 394
FT DOMAIN 45 126
FT DOMAIN 142 394
FT ACT_SITE 189 189
FT ACT_SITE 238 238
FT ACT_SITE 345 345
FT DISULFID 45 126
FT DISULFID 66 108
FT DISULFID 97 121
FT DISULFID 131 262
FT DISULFID 174 190
FT DISULFID 182 251
FT DISULFID 276 351
FT DISULFID 308 324
FT DISULFID 341 369
FT CARBOHYD 315 315
SQ SEQUENCE 394 AA; 44105 MW; 9CCD6F52F3D81FCD CRC64;

Query Match 36.2%; Score 546; DB 1; Length 394;
Best Local Similarity 42.9%; Pred. No. 5.9e-42;
Matches 118; Conservative 43; Mismatches 102; Indels 12; Gaps 6

Qy 4 SPPEELKFCQG-QKTLRPRFKIIIGGEFTTIENQWFAAIYRRHRGGS-VTVCGGSLISP 61
Db 122 SVFVCSKATCGLRKYKEPQLHSTGGLFTDITSHFWQAIPAQRSSSGERLCCGILLSS 181
Qy 62 CWTISATHCFID-IPKEDYIVYLGSRNLNSNTGEMKFEVENTILHKDVSADTLAHND 120
Db 182 CWLTAHCFQERYPPQHLRVV-LGRYVRKPGKEQTFFVEKCIIEHEFDDDT--YNNND 238
Qy 121 IALLKIRSKERCAQPSRTIOTICLPMSYNDPQGTSCETIGKGNSTDYLYPEOLKMT 180
Db 239 IALLQKSGSPQCAQESDSVRAICLPEANLQLPDWTCELSGKGKSSFFSEQLKEG 298
Qy 181 VKLISHRECOQPHYGSEVTTKMLCAADPOWKT-----DSCQDGGGGLVCSLQGRMT 234
Db 239 HVRYPSSRCTSKFLNKTNTNNMLCAGTRSGBEYIPNVHDAQCQDGGGGLVCMNDNMT 358
Qy 235 LTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTK 269
Db 359 LGGIISWGVGGEKDFGVYTKVNYLGIWIRDNR 393

RESULT 14
ID URT2 DESRO STANDARD; PRT; 477 AA.
AC P15638;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1996 (Rel. 33, last sequence update)
DE 28-FEB-2003 (Rel. 41, last annotation update)
DE Salivary plasminogen activator alpha 2 precursor (EC 3.4.21.68) (DSPA alpha-2) (BAT-PA) (T-plasminogen activator).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
```



RP TISSUE-Salivary Gland;  
RC MEDLINE=92039036; PubMed=1937019;  
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,  
Alagon A., Donner P., Schleuning W.D.;  
RT "The plasminogen activator family from the salivary gland of the  
vampire bat *Desmodus rotundus*: cloning and expression.";  
RL Gene 105:229-237(1991).  
RN [2]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE-Salivary Gland;  
RX MEDLINE=93036867; PubMed=2509450;  
RA Gardell S.J., Duong L.T., Diehl R.E., York J.D., Hare T.R.,  
Register R.B., Jacobs J.W., Dixon R.A.F., Friedman P.A.;  
RT "Isolation, characterization, and cDNA cloning of a vampire bat  
salivary plasminogen activator";  
RL J. Biol. Chem. 264:17947-17952(1989).  
RN [3]  
RP CHARACTERIZATION.  
RX MEDLINE=9339059; PubMed=1309059;  
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,  
Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,  
Donner P.;  
RT "Plasminogen activators from the saliva of *Desmodus rotundus* (common  
vampire bat): unique fibrin specificity";  
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).  
CC -!- FUNCTION: Probably essential to support the feeding habits of this  
exclusively haematophagous animal. Probable potent thrombolytic  
agent.  
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
plasminogen to form plasmin.  
CC -!- ENZYME REGULATION: Activity toward plasminogen is stimulated in  
the presence of fibrin I.  
CC -!- SUBUNIT: Monomer.  
CC -!- DOMAIN: The fibronectin type-I domain mediates binding to fibrin,  
and the kringle domain apparently mediates fibrin-induced  
stimulation of activity.  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
CC -!- SIMILARITY: Contains 1 EGF-like domain.  
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.  
CC -!- SIMILARITY: Contains 1 kringle domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; M63988; AAA31593.1; -;  
DR EMBL; J05082; AAA31596.1; -;  
DR PIR; A34369; A34369.  
DR HSSP; P98119; LA51.  
DR MEROPS; S01.232; -;  
DR InterPro; IPR009003; Cys Ser trypsin.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR000083; Fibronctn.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR001254; Peptidase S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00008; EGF; 1.  
DR Pfam; PF00039; fn1; 1.  
DR Pfam; PF00051; kringle; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00181; EGF; 1.  
DR SMART; SM00058; FN1; 1.  
DR SMART; SM00130; KR; 1.

DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS01186; EGF\_2; 1.  
DR PROSITE; PS00026; EGF\_3; 1.  
DR PROSITE; PS01253; FIBRONECTIN 1; 1.  
DR PROSITE; PS00021; KRINGLE 1; 1.  
DR PROSITE; PS00070; KRINGLE 2; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
DR PROSITE; PS00136; TRYPSIN\_HIS; 1.  
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
KW Kringle; EGF-like domain; Signal; Multigene family.  
FT SIGNAL 1 36 POTENTIAL.  
FT CHAIN 37 477 SALIVARY PLASMINOGEN ACTIVATOR ALPHA 2.  
FT DOMAIN 40 82 FIBRONECTIN TYPE-1.  
FT DOMAIN 83 121 EGF-LIKE.  
FT DOMAIN 128 209 KRINGLE.  
FT DOMAIN 225 477 SERINE PROTEASE.  
FT ACT\_SITE 272 272 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 321 321 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 428 428 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT DISULFID 42 72 BY SIMILARITY.  
FT DISULFID 70 79 BY SIMILARITY.  
FT DISULFID 87 98 BY SIMILARITY.  
FT DISULFID 92 109 BY SIMILARITY.  
FT DISULFID 111 120 BY SIMILARITY.  
FT DISULFID 128 209 BY SIMILARITY.  
FT DISULFID 149 191 BY SIMILARITY.  
FT DISULFID 180 204 BY SIMILARITY.  
FT DISULFID 214 345 BY SIMILARITY.  
FT DISULFID 257 273 BY SIMILARITY.  
FT DISULFID 265 334 BY SIMILARITY.  
FT DISULFID 359 434 BY SIMILARITY.  
FT DISULFID 391 407 BY SIMILARITY.  
FT DISULFID 424 452 BY SIMILARITY.  
FT CARBOHYD 185 185 N-LINKED (GLCNAC... ) (POTENTIAL).  
FT CARBOHYD 398 398 N-LINKED (GLCNAC... ) (POTENTIAL).  
FT CONFLICT 403 403 N -> K (IN REF. 2).  
FT CONFLICT 417 417 Y -> H (IN REF. 2).  
FT CONFLICT 435 435 M -> R (IN REF. 2).  
SQ SEQUENCE 477 AA; 53719 MW; 17486555C0E5077C CRC64;  
Query Match 36.2%; Score 546; DB 1; Length 477;  
Best Local Similarity 42.9%; Pred No. 7.5e-42;  
Matches 118; Conservative 43; Mismatches 102; Indels 12; Gaps 6  
Qy 4 SPPEELKFCQG-QKTLRPRFKIIGGFTTIENOFWFAIYRHRGGS-VYVCGSLISP 61  
Db 205 SVPVCSKATCLGRKYKEPOLHSTGGLFTDITSHPOAAIPQNRSSGRFLCGGILISS 264  
Qy 62 CWVTSATHCTID-YPKKEDYIVVGLSRSLNSNTQGMKREVENLILHKDYSADTLAHND 120  
Db 265 CWVLTAAHCFQERYPPHOLRVV-LGRYRVKPKREQTFEVKCIHVEEFDDET--YND 321  
Qy 121 IALLKIRSKRCRCAQPSRTIQTCLPMSYNDPQGTSCETGFGKENSTDYLYPEQLKMT 180  
Db 322 IALLQLKSGSPQCAQESDSVRAICLPEANLQLPDWTECELSGYGKHKSSPPFYSEQLKEG 381  
Qy 181 VKVLISHRECCQPHYGYGSEVTTOMLCAADPQWKT-----DSCGGSGGLVCSLQGRMT 234  
Db 382 HVRLYPSSRCTSKFLFNKVTNNMLCAGDTRSGEIVFNVDHACQGGSGGLVCSLQGRMT 441  
Qy 235 LTGIVSWGRGKALKDKPQGVYTRVSHPLPMTIRSHTK 269  
Db 442 LGLIISWGVGGEKIDPGVYTKVNYLGRDNR 476  
RESULT 15  
ID UR11 DESRO STANDARD; PRT; 477 AA.  
AC P98119; AC  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Salivary plasminogen activator alpha 1 precursor (BC 3.4.21.68) (DSFA  
DE alpha-1).  
OS Desmodus rotundus (Vampire bat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;  
OC Desmodontinae; Desmodus.  
OX NCBI\_TaxID=9430;  
RN [1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Salivary Gland;  
RX MEDLINE=92019036; PubMed=1937019;  
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,  
RA Alagon A., Donner P., Schlenning W.D.;  
RT "The plasminogen activator family from the salivary gland of the  
RL vampire bat Desmodus rotundus: cloning and expression.";  
RL Gene 105:229-237(1991).  
RN [2]  
RP CHARACTERIZATION.  
RX MEDLINE=93393059; PubMed=1309059;  
RA Schlenning W.D., Alagon A., Boidol W., Bringmann P., Petri T.,  
RA Kraetzschmar J., Haendler B., Langer G., Baidus B., Witt W.,  
RA Donner P.;  
RT "Plasminogen activators from the saliva of Desmodus rotundus (common  
RT vampire bat): unique fibrin specificity.";  
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).  
RN [3]  
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
RC TISSUE=salivary gland;  
RX MEDLINE=98022741; PubMed=9354616;  
RA Renatus M., Stubbs M.T., Huber R., Bringmann P., Donner P.,  
RA Schlenning W.D., Bode W.;  
RT "Catalytic domain structure of vampire bat plasminogen activator: a  
RT molecular paradigm for proteolysis without activation cleavage.";  
RL Biochemistry 36:13483-13493(1997).  
CC -!- FUNCTION: Probably essential to support the feeding habits of this  
CC exclusively haematophagous animal. Potent thrombolytic agent.  
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
CC plasminogen to form plasmin.  
CC -!- ENZYME REGULATION: Activity toward plasminogen is stimulated in  
CC the presence of fibrin I.  
CC -!- SUBUNIT: Monomer.  
CC -!- DOMAIN: The fibronectin type-I domain mediates binding to fibrin,  
CC and the kringle domain apparently mediates fibrin-induced  
CC stimulation of activity.  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
CC -!- SIMILARITY: Contains 1 EGF-like domain.  
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.  
CC -!- SIMILARITY: Contains 1 kringle domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M63987; AAA31591.1; -;  
CC EMBL; M63986; AAA31592.1; -;  
CC PIR; JS0597; JS0597.  
CC PDB; 1A5I; 23-MAR-99.  
CC MEROPS; S01.232; -;  
CC GlycoSuiteDB; P98119; -;  
CC InterPro; IPR003003; Cys\_Ser\_trypsin.  
CC InterPro; IPR006209; EGF\_like.  
CC InterPro; IPR000083; Fibrinctnl.  
CC InterPro; IPR006210; IEFG.  
CC InterPro; IPR000001; Kringle.  
CC InterPro; IPR001254; Peptidase\_S1.  
CC InterPro; IPR001314; Peptidase\_S1A.  
CC Pfam; PF00008; EGF; 1.  
CC Pfam; PF00039; Fni; 1.

DR Pfam; PF00051; kringle; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00181; EGF; 1.  
DR SMART; SM00058; FNI; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS01186; EGF\_2; 1.  
DR PROSITE; PS50026; EGF\_3; 1.  
DR PROSITE; PS01253; FIBRONECTIN\_1; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS00070; KRINGLE\_2; 1.  
DR PROSITE; PS50240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
KW Kringle; EGF-like domain; Signal; Multi-gene family; 3D-structure.  
FT SIGNAL 1 36 POTENTIAL.  
FT CHAIN 37 477 SALIVARY PLASMINOGEN ACTIVATOR ALPHA 1.  
FT DOMAIN 40 82 FIBRONECTIN TYPE-I.  
FT DOMAIN 83 121 EGF-LIKE.  
FT DOMAIN 128 209 KRINGLE.  
FT DOMAIN 225 477 SERINE PROTEASE.  
FT ACT\_SITE 272 272 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 321 321 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 428 428 CHARGE RELAY SYSTEM.  
FT DISULFID 42 72 BY SIMILARITY.  
FT DISULFID 70 79 BY SIMILARITY.  
FT DISULFID 87 98 BY SIMILARITY.  
FT DISULFID 92 109 BY SIMILARITY.  
FT DISULFID 111 120 BY SIMILARITY.  
FT DISULFID 128 209 BY SIMILARITY.  
FT DISULFID 149 191 BY SIMILARITY.  
FT DISULFID 180 204 BY SIMILARITY.  
FT DISULFID 214 345 BY SIMILARITY.  
FT DISULFID 257 273 BY SIMILARITY.  
FT DISULFID 265 334 BY SIMILARITY.  
FT DISULFID 359 434 BY SIMILARITY.  
FT DISULFID 391 407 BY SIMILARITY.  
FT DISULFID 424 452 BY SIMILARITY.  
FT CARBOHYD 153 153  
FT CARBOHYD 398 398  
FT TURN 214 215  
FT STRAND 223 224  
FT TURN 226 227  
FT STRAND 230 231  
FT HELIX 234 236  
FT TURN 238 239  
FT STRAND 240 245  
FT STRAND 254 263  
FT TURN 264 265  
FT STRAND 266 269  
FT HELIX 271 273  
FT TURN 280 282  
FT STRAND 284 287  
FT TURN 297 298  
FT STRAND 300 309  
FT TURN 311 312  
FT TURN 315 317  
FT TURN 319 320  
FT STRAND 323 328  
FT STRAND 338 338  
FT TURN 339 340  
FT STRAND 341 341  
FT STRAND 345 345  
FT TURN 349 350  
FT TURN 355 356  
FT STRAND 358 363

N-LINKED (GLCNAC. . .).  
/FTid=CAR\_000027.  
N-LINKED (GLCNAC. . .).  
/FTid=CAR\_000028.

FT	STRAND	366	366	
FT	STRAND	374	374	
FT	STRAND	379	385	
FT	HELIX	388	390	
FT	TURN	393	398	
FT	TURN	403	404	
FT	STRAND	405	409	
FT	TURN	425	426	
FT	TURN	428	439	
FT	STRAND	431	436	
FT	TURN	437	438	
FT	STRAND	439	448	
FT	TURN	455	456	
FT	STRAND	459	463	
FT	HELIX	464	467	
FT	HELIX	468	474	
SQ	SEQUENCE	477 AA; 53616 MW; AA06FD1739C10E5E CRC64;		
Query Match				
Best Local Similarity 36.0%; Score 543; DB 1; Length 477;				
Matches 117; Conservative 42; Mismatches 98; Indels 20; Gaps 6;				
QY	4	SPPEELKFCG-QKTLRPRFKIIGGEFTTIENQPFPAAYRRHRGGS-VTVVCGGSLISP	61	
DB	205	SVPVCSKATCGLRKYKEPOLHSTGGLFTDITSHPWQAALPAQRSSGERFLCGGILLIS	264	
QY	62	CWVISATHCFIDYPPKEDYI-----VYGRSLNNTQGMKFEVENILHDKYSADTLA	116	
DB	265	CWVLTAAHCF-----QESYLPQLKVLGRVYRKPGEETFKYKVIHVHKEPDDP--	317	
QY	117	HNNDIALKIRKSGRCAPSRITQICLPMSYNDPQGTSCETGFGKSTNDLYLPEQ	176	
DB	318	YNNDIALLQKSDSPCAESDSVRAICLPEANLQLPDWTELGSGYKHKSSPFYSEQ	377	
QY	177	LQNTVVKLISHRECQOPHYGSEVTTKMLCAADPQWKT-----DSCQDGGGPIVCSLQ	230	
DB	378	LKEGHVRLYPSRSCAPKFLFNKTVTNNMLCAGDTSGEIYPNVHDACQDGGGPIVCMND	437	
QY	231	GRMTLGIYSWGRCALKDKPGVYTRVSHFLWIRSH	267	
DB	438	NHMTLLGIISWGVCCKEDKDPGVYTKVNYLGIWRDN	474	
RESULT 16				
HGFA MOUSE				
ID	HGFA_MOUSE	STANDARD;	PRT;	653 AA.
AC	Q9R028; Q9JKV4;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF			
DE	activator) (HGFA).			
GN	HGFAC.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BALB/c;			
RA	Itch H., Kataoka H., Koono H.;			
RT	"Mouse hepatocyte growth factor activator.";			
RL	Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.			
RL	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21226753; PubMed=11032833;			
RA	van Adelsberg J.S., Sehgal S., Kukes A., Brady C., Barasch J.,			
RA	Yang J., Huan Y.;			
RT	"Activation of hepatocyte growth factor (HGF) by endogenous HGF			
RT	activator is required for mesangial kidney morphogenesis in			
RT	vitro.";			
RL	J. Biol. Chem. 276:15099-15106(2001).			
CC	-!- FUNCTION: Activates hepatocyte growth factor (HGF) by converting			

CC	it from a single chain to a heterodimeric form (By similarity).			
CC	-!- SUBUNIT: Dimer of a short chain and a long chain linked by a			
CC	disulfide bond (By similarity).			
CC	-!- SUBCELLULAR LOCATION: Secreted as an inactive single-chain			
CC	precursor and is then activated to a heterodimeric form (By			
CC	similarity).			
CC	-!- SIMILARITY: Belongs to peptidase family S1.			
CC	-!- SIMILARITY: Contains 2 EGF-like domains.			
CC	-!- SIMILARITY: Contains 1 fibronectin type I domain.			
CC	-!- SIMILARITY: Contains 1 fibronectin type II domain.			
CC	-!- SIMILARITY: Contains 1 kringle domain.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercia			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce">http://www.isb-sib.ch/announce</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
CC	EMBL; AF099017; AAF02489.1; -			
DR	EMBL; AF224724; AAF34712.1; -			
DR	HSP; P00763; IDPO.			
DR	MEROPS; S01.228; -			
DR	MGD; MGI:1859281; Hgfac.			
DR	InterPro; IPR009003; Cys_ser_trypsin.			
DR	InterPro; IPR000742; EGF_2			
DR	InterPro; IPR006209; EGF-like.			
DR	InterPro; IPR000083; Fibrinctni.			
DR	InterPro; IPR000562; FN_Type_II.			
DR	InterPro; IPR006210; IEGF.			
DR	InterPro; IPR000001; Kringle.			
DR	InterPro; IPR001254; Peptidase_S1.			
DR	InterPro; IPR001314; Peptidase_S1A.			
DR	Pfam; PF00008; EGF; 2.			
DR	Pfam; PF00039; fn1; 1.			
DR	Pfam; PF00040; fn2; 1.			
DR	Pfam; PF00051; kringle; 1.			
DR	Pfam; PF00089; trypsin; 1.			
DR	PRINTS; PR00722; CHYMOTRYPSIN.			
DR	PRINTS; PR00013; FNTYPEII.			
DR	PRINTS; PR00018; KRINGLE.			
DR	ProDom; PD000995; FN_Type_II; 1.			
DR	ProDom; PD000395; Kringle; 1.			
DR	SMART; SM00181; EGF; 2.			
DR	SMART; SM00059; FN2; 1.			
DR	SMART; SM00130; KR; 1.			
DR	SMART; SM00020; Tryp_SPC; 1.			
DR	PROSITE; PS00022; EGF_1; 2.			
DR	PROSITE; PS01186; EGF_2; 1.			
DR	PROSITE; PS00026; EGF_3; 2.			
DR	PROSITE; PS01253; FIBRONECTIN_1; 1.			
DR	PROSITE; PS00023; FIBRONECTIN_2; 1.			
DR	PROSITE; PS00021; KRINGLE_1; 1.			
DR	PROSITE; PS00070; KRINGLE_2; 1.			
DR	PROSITE; PS00240; TRYPSIN_DOM; 1.			
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.			
DR	PROSITE; PS00135; TRYPSIN_SER; 1.			
KW	Hydrolase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;			
KW	EGF-like domain; Repeat; Zymogen.			
FT	SIGNAL	1	29	BY SIMILARITY.
FT	PROPEP	30	369	CLEAVED IN ACTIVE FORM (BY SIMILARITY).
FT	CHAIN	370	405	HEPATOCYTE GROWTH FACTOR ACTIVATOR SHORT
FT	CHAIN			CHAIN.
FT	CHAIN	406	653	HEPATOCYTE GROWTH FACTOR ACTIVATOR LONG
FT	CHAIN			CHAIN.
FT	DOMAIN	105	145	FIBRONECTIN TYPE-II.
FT	DOMAIN	157	195	EGF-LIKE 1.
FT	DOMAIN	197	237	FIBRONECTIN TYPE-I.
FT	DOMAIN	238	276	EGF-LIKE 2.
FT	DOMAIN	283	364	KRINGLE.
FT	DOMAIN	406	653	SERINE PROTEASE.
FT	ACT_SITE	445	445	CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT SITE 495 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT SITE 596 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT DISULFID 105 BY SIMILARITY.

FT DISULFID 119 BY SIMILARITY.

FT DISULFID 145 BY SIMILARITY.

FT DISULFID 161 BY SIMILARITY.

FT DISULFID 183 BY SIMILARITY.

FT DISULFID 194 BY SIMILARITY.

FT DISULFID 227 BY SIMILARITY.

FT DISULFID 234 BY SIMILARITY.

FT DISULFID 242 BY SIMILARITY.

FT DISULFID 253 BY SIMILARITY.

FT DISULFID 264 BY SIMILARITY.

FT DISULFID 275 BY SIMILARITY.

FT DISULFID 283 BY SIMILARITY.

FT DISULFID 304 BY SIMILARITY.

FT DISULFID 335 BY SIMILARITY.

FT DISULFID 392 BY SIMILARITY.

FT DISULFID 430 BY SIMILARITY.

FT DISULFID 438 BY SIMILARITY.

FT DISULFID 533 BY SIMILARITY.

FT DISULFID 565 BY SIMILARITY.

FT DISULFID 592 BY SIMILARITY.

FT CARBOHYD 39 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 47 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 63 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 287 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 466 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 544 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CONFLICT 164 G -> W (IN REF. 2).

SQ SEQUENCE 653 AA; 70567 MW; 88B4B20255DF7FDC CRC64;

Query Match 33.0%; Score 497.5; DB 1; Length 653;

Best Local Similarity 38.5%; Pred. No. 2.7e-37;

Matches 105; Conservative 44; Mismatches 109; Indels 15; Gaps 6;

QY 2 PSSPPELKPQCOOK-----TLPRPKLIGGETTIENOPWPAAYRRHSGSVTVVCGG 56

Db 382 PESAP-AVRPTCGKRRKRTFLPR--IIGSSSLPGSHPLAAIY---IGNS---FCAG 432

QY 57 SLISPCWVISAATHCFIDYPKEDYIVVGLGRSLNSNTQGMKPEVENLHKDYSADTLA 116

Db 433 SLVHTCWVSAACHFANSPRDSITVVLGQFFNRITDTQIFGIEKYVPTLYSVNFEN 492

QY 117 HNDIALKIRSEKRCQPSRTQITCLPSMNDPQFSGTSCITGFKENSTDYLYPEQ 176

Db 493 NH-DLVILRLKXKGERCAVRSQFVQPCILPEAGSFPTGKQIAGMGHMDENVSSYSN 551

QY 177 LKMTVVKLISHRECOQPHYGVSEVTTQMLCAADPQWKTSCQDGGPLVCSLQGRMTLT 236

Db 552 LLEALVPLVADHKCSPEVYGDALSPNNLCAGYFCDKSDACQDGGPLVCEKNGVAYLY 611

QY 237 GIVSWGRCALKKXPGVYTVRSHPLPWRSHTK 269

Db 612 GIISWGDGGRINKPGVYTVRANYVDWINDRIR 644

RESULT 17

FA12 HUMAN

ID FA12\_HUMAN STANDARD; PRT; 615 AA.

AC P00748; P78339;

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-OCT-1989 (Rel. 12, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)

GN F12.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=86007593; PubMed=2868762;

RA Cool D.E., McGillivray R.T.A.;

"Characterization of the human blood coagulation factor XII gene. Intron/exon gene organization and analysis of the 5'-flanking region."; J. Biol. Chem. 262:13662-13673 (1987).

[2]

RP SEQUENCE FROM N.A., AND VARIANTS ALA-207; ASP-545 AND HIS-605. Rieder M.J., Amel T.Z., Carrington D.P., Ozuna M., Kuldane S.A., Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.; Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.

[3]

RP SEQUENCE OF 4-615 FROM N.A. MEDLINE=86176794; PubMed=3754331; Tripodi M., Citarella F., Guida S., Galeffi P., Fantoni A., Cortese R.; "cDNA sequence coding for human coagulation factor XII (Hageman)."; Nucleic Acids Res. 14:1346-1346 (1986).

[4]

RP SEQUENCE OF 14-615 FROM N.A. MEDLINE=86033830; PubMed=3877053; Cool D.E., Edgell C.-J.S., Louie G.V., Zoller M.J., Brayer G.D., McGillivray R.T.A.; "Characterization of human blood coagulation factor XII cDNA. Prediction of the primary structure of factor XII and the tertiary structure of beta-factor XIIa."; J. Biol. Chem. 260:13666-13676 (1985).

[5]

RP SEQUENCE OF 146-615 FROM N.A. MEDLINE=86216049; PubMed=3011063; Que B.G., Davie E.W.; "Characterization of a cDNA coding for human factor XII (Hageman factor)."; Biochemistry 25:1525-1528 (1986).

[6]

RP SEQUENCE OF 20-379. MEDLINE=85182674; PubMed=3886654; McMullen B.A., Fujikawa K.; "Amino acid sequence of the heavy chain of human alpha-factor XIIa (activated Hageman factor)."; J. Biol. Chem. 260:5328-5341 (1985).

[7]

RP SEQUENCE OF 354-362 AND 373-615. MEDLINE=83291041; PubMed=6604055; Fujikawa K., McMullen B.A.; "Amino acid sequence of human beta-factor XIIa."; J. Biol. Chem. 258:10924-10933 (1983).

[8]

RP SEQUENCE OF 561-615 FROM N.A. TISSUE=Blood; MEDLINE=96133302; PubMed=8528215; Schloesser M., Hofferbert S., Bartz U., Lutze G., Lammle B., Engel W.; "The novel acceptor splice site mutation 11396(G-->A) in the factor XII gene causes a truncated transcript in cross-reacting material negative patients."; Hum. Mol. Genet. 4:1235-1237 (1995).

[9]

RP CARBOHYDRATE-LINKAGE SITE THR-109. MEDLINE=92184750; PubMed=1544894; Harris R.J., Ling V.T., Spellman M.W.; "O-linked fucose is present in the first epidermal growth factor domain of factor XII but not protein C."; J. Biol. Chem. 267:5102-5107 (1992).

[10]

RP VARIANT WASHINGTON D.C. SER-590. MEDLINE=90046788; PubMed=2510163; Miyata T., Kawabata S.-I., Iwanaga S., Takahashi I., Alving B., Saito H.; "Coagulation factor XII (Hageman factor) Washington D.C.: inactive factor XIIa results from Cys-571-->Ser substitution."; Proc. Natl. Acad. Sci. U.S.A. 86:8319-8322 (1989).

[11]

RP VARIANT LOCARNO PRO-372. MEDLINE=94325559; PubMed=8049433; Hovinga J.K., Schaller J., Stricker H., Willemin W.A., Furlan M.,

Laemmle B.;  
 "Coagulation factor XII Locarno: the functional defect is caused by the amino acid substitution Arg-353-->Pro leading to loss of a kallikrein cleavage site."  
 Blood 84:1173-1181(1994).  
 [12]  
 RP VARIANT TENRI CYS-53.  
 RX MEDLINE=99290785; PubMed=10361128;  
 RA Kondo S., Tokunaga F., Kawano S., Oono Y., Kumagai S., Koide T.;  
 RT "Factor XII Tenri, a novel cross-reacting material negative factor XII deficiency, occurs through a proteasome-mediated degradation."  
 RL Blood 93:4300-4308(1999).  
 CC -!- FUNCTION: Factor XII is a serum glycoprotein that participates in the initiation of blood coagulation, fibrinolysis, and the generation of bradykinin and angiotensin.  
 CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Ile bonds in factor VII to form factor VIIIa and factor XI to form factor Xla.  
 CC -!- PTM: O- AND N-GLYCOSYLATED.  
 CC -!- DISEASE: Defects in F12 do not cause any clinical symptoms. The sole effect is that whole-blood clotting time is prolonged.  
 CC -!- MISCELLANEOUS: Factor XII, prekallikrein, and HMW kininogen form a complex bound to an anionic surface. Prekallikrein is cleaved by factor XII to form kallikrein, which then cleaves factor XII first to alpha-factor Xla and then to beta-factor Xlaa. Alpha-factor Xlaa activates factor XI to factor Xla.  
 CC -!- SIMILARITY: Belongs to peptidase family S1.  
 CC -!- SIMILARITY: Contains 2 EGF-like domains.  
 CC -!- SIMILARITY: Contains 1 fibronectin type I domain.  
 CC -!- SIMILARITY: Contains 1 fibronectin type II domain.  
 CC -!- SIMILARITY: Contains 1 kringle domain.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; M31315; AAA70225.1; -;  
 DR EMBL; AF538691; AAM97932.1; -;  
 DR EMBL; M11723; AAA51986.1; -;  
 DR EMBL; M17465; AAB59490.1; -;  
 DR EMBL; M17465; AAB59490.1; JOINED.  
 DR EMBL; M17465; AAB59490.1; JOINED.  
 DR EMBL; M13147; AAA70224.1; -;  
 DR EMBL; U71274; AAB51203.1; -;  
 DR PIR; A29411; KFHU12.  
 DR HSSP; P00763; IDPO.  
 DR MROPS; S01.211; -;  
 DR Genew; HGNC:3530; F12.  
 DR MIM; 234000; -;  
 DR GO; GO:0003805; F:blood coagulation factor XI activity; TAS.  
 DR GO; GO:0003806; F:blood coagulation factor XII activity; TAS.  
 DR GO; GO:0008236; F:serine-type peptidase activity; TAS.  
 DR GO; GO:0007596; P:blood coagulation; TAS.  
 DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
 DR InterPro; IPR006209; EGF\_like.  
 DR InterPro; IPR000083; Fibrinctnl.  
 DR InterPro; IPR000562; FN\_Type\_II.  
 DR InterPro; IPR006210; IEGF.  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR Pfam; PF00008; EGF; 2.  
 DR Pfam; PF00039; fn1; 1.  
 DR Pfam; PF00040; fn2; 1.  
 DR Pfam; PF00051; kringle; 1.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PRINTS; PR00013; FNTYPEII.  
 DR PRINTS; PR00018; KRINGLE.  
 DR ProDom; PD000995; FN\_Type\_II; 1.

ProDom; PD000395; Kringle; 1.  
 DR SMART; SM00181; EGF; 2.  
 DR SMART; SM00058; FN1; 1.  
 DR SMART; SM00059; FN2; 1.  
 DR SMART; SM00130; KR; 1.  
 DR SMART; SM00020; Tryp\_Spc; 1.  
 DR PROSITE; PS00022; EGF\_1; 2.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS00026; EGF\_3; 2.  
 DR PROSITE; PS01253; FIBRONECTIN\_1; 1.  
 DR PROSITE; PS00023; FIBRONECTIN\_2; 1.  
 DR PROSITE; PS00021; KRINGLE\_1; 1.  
 DR PROSITE; PS00070; KRINGLE\_2; 1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;  
 KW Hydrolase; Fibrinolysis; Signal; EGF-like domain; Repeat; Zymogen;  
 KW Polymorphism; Disease mutation.  
 FT SIGNAL 1 19 ALPHA-FACTOR X1A HEAVY CHAIN.  
 FT CHAIN 20 372 ALPHA-FACTOR X1A LIGHT CHAIN.  
 FT CHAIN 373 615 BETA-FACTOR X1A PART 1.  
 FT CHAIN 354 362 BETA-FACTOR X1A PART 2.  
 FT CHAIN 373 615 FIBRONECTIN TYPE-II.  
 FT DOMAIN 47 88 FIBRONECTIN TYPE-I.  
 FT DOMAIN 94 131 EGF-LIKE 1.  
 FT DOMAIN 133 173 FIBRONECTIN TYPE-I.  
 FT DOMAIN 174 210 EGF-LIKE 2.  
 FT DOMAIN 217 295 KRINGLE.  
 FT DOMAIN 296 349 PRO-RICH.  
 FT DOMAIN 373 615 SERINE PROTEASE.  
 FT CARBOHYD 109 109 O-LINKED (FUC).  
 FT CARBOHYD 249 249 N-LINKED (GLCNAC...).  
 FT CARBOHYD 299 299 O-LINKED (POTENTIAL).  
 FT CARBOHYD 305 305 O-LINKED (POTENTIAL).  
 FT CARBOHYD 308 308 O-LINKED (POTENTIAL).  
 FT CARBOHYD 328 328 O-LINKED (POTENTIAL).  
 FT CARBOHYD 329 329 O-LINKED (POTENTIAL).  
 Query Match 32.8%; Score 495; DB 1; Length 615;  
 Best Local Similarity 38.5%; Pred. No. 4.3e-37;  
 Matches 107; Conservative 44; Mismatches 109; Indels 18; Gaps 6  
 Qy 1 KPSSPPBELK---PQCQ---KTLRPFKLIIGSEFTIENQPFALYRHGGSVTVVC 54.  
 Db 344 KREQPFSLTRNGPLSCGRLKLSLSSMTVVGVGLVARGAHPYALYWGHS-----FC 397  
 Qy 55 GGSLSIPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENILHKQYSADT 114  
 Db 398 AGSLIAPCWVLTAAHCLQDRPAPEDLTVLGQERRNHSCEPQTALVRSYRLHEAPS--P 455  
 Qy 115 LAHNDIALKIR-SKEGRCAQPSRTTQTLCLPSMYNDPQPGTSCEITGPKENSTDYLY 173  
 Db 456 VSYQDLALLRLQEDADGSCALLSPYQVCLPSGAARPSETTLCQVAGWHQFEGAEY 515  
 Qy 174 PEQLKMTVVKLISHRECOQPHYGVSEVTKMLCAADPQWKTDSCQGDGGLPLVCSLQG-- 231  
 Db 516 ASFLQEAQVPEFLSLERCSAPDVHGSSILPGMLCAGFLEGGTDACQGDGGLVCEQAAE 575  
 Qy 232 -RMLTGVSWGRGALKKDPGVTVTRVSHFLPWIRSH 268  
 Db 576 RRLTLQGIISWGGCGDRNKPEGVYTDVAYLAWIREHT 613  
 RESULT 18  
 HGFA HUMAN  
 ID HGFA HUMAN STANDARD; PRT; 655 AA.  
 AC Q04756; Q14726;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF activator) (HGFA).

GN HGFAC.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RP TISSUE=Liver, and Serum;  
RX MEDLINE=93252878; PubMed=7693665;  
RA Miyazawa K., Shimomura T., Kitamura A., Kondo J., Morimoto Y.,  
RA Kitamura N.;  
RT "Molecular cloning and sequence analysis of the cDNA for a human  
RT serine protease responsible for activation of hepatocyte growth  
RT factor. Structural similarity of the protease precursor to blood  
RT coagulation factor XII.";  
RL J. Biol. Chem. 268:10024-10028(1993).  
RN [2]  
RN SEQUENCE OF 40-655 FROM N.A.  
RA Zhao S., Odell C.;  
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Activates hepatocyte growth factor (HGF) by  
CC converting it from a single chain to a heterodimeric form.  
CC -!- SUBUNIT: Dimer of a short chain and a long chain linked by a  
CC disulfide bond.  
CC -!- SUBCELLULAR LOCATION: Secreted as an inactive single-chain  
CC precursor and is then activated to a heterodimeric form.  
CC -!- TISSUE SPECIFICITY: Liver.  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
CC -!- SIMILARITY: Contains 2 EGF-like domains.  
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.  
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.  
CC -!- SIMILARITY: Contains 1 kringle domain.  
CC -!- CAUTION: It is uncertain whether Met-1 is the initiator.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; D14012; BAA03113.1; -;  
DR PIR; Z69923; -; NOT\_ANNOTATED\_CDS.  
DR PIR; A46688; A46688.  
DR HSSP; P00763; IDPO.  
DR MEROPS; S01.228; -;  
DR Genew; HGNC:4894; HGFAC.  
DR MIM; 604552; -;  
DR GO; GO:0005576; C:extracellular; TAS.  
DR GO; GO:0004252; F:serine-type endopeptidase activity; TAS.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; TAS.  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR000742; EGF\_2.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR000083; Fibnctnl.  
DR InterPro; IPR000562; FN\_type\_I.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00008; EGF; 2.  
DR Pfam; PF00039; fn1; 1.  
DR Pfam; PF00040; fn2; 1.  
DR Pfam; PF00051; kringle; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00013; FNTYPEII.  
DR PRINTS; PR00018; FNTYPEII.  
DR ProDom; PD000995; FN\_Type\_II; 1.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00181; EGF; 2.  
DR SMART; SM00058; FN1; 1.

DR SMART; SMC0059; FN2; 1.  
DR SMART; SMC0130; KR; 1.  
DR SMART; SMC0020; TYP\_SPC; 1.  
DR PROSITE; PS00022; EGF\_1; 2.  
DR PROSITE; PS01186; EGF\_2; 1.  
DR PROSITE; PS00026; EGF\_3; 2.  
DR PROSITE; PS01253; FIBRONECTIN\_1; 1.  
DR PROSITE; PS00023; FIBRONECTIN\_2; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS00070; KRINGLE\_2; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolyase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;  
KW EGF-like domain; Repeat; Zymogen.  
FT SIGNAL 1 30  
FT PROPEP 31 372  
FT CHAIN 373 407  
FT CHAIN 408 655  
FT CHAIN 108 148  
FT DOMAIN 160 198  
FT DOMAIN 200 240  
FT DOMAIN 241 279  
FT DOMAIN 286 367  
FT DOMAIN 408 655  
FT ACT\_SITE 447 447  
FT ACT\_SITE 497 497  
FT ACT\_SITE 598 598  
FT DISULFID 108 133  
FT DISULFID 122 148  
FT DISULFID 164 175  
FT DISULFID 169 186  
FT DISULFID 188 197  
FT DISULFID 202 230  
FT DISULFID 228 237  
FT DISULFID 245 256  
FT DISULFID 250 267  
FT DISULFID 269 278  
FT DISULFID 286 367  
FT DISULFID 307 349  
FT DISULFID 338 362  
FT DISULFID 394 521  
FT DISULFID 432 448  
FT DISULFID 440 510  
FT DISULFID 535 604  
FT DISULFID 567 583  
FT DISULFID 594 622  
FT CARBOHYD 48 48  
FT CARBOHYD 290 290  
FT CARBOHYD 468 468  
FT CARBOHYD 492 492  
FT CARBOHYD 546 546  
FT CONFLICT 644 644  
SQ SEQUENCE 655 AA; 70681 MW; 2CF72F1E1B862ED7 CRC64;  
Query Match 32.7%; Score 493.5; DB 1; Length 655;  
Best Local Similarity 38.3%; Pred. No. 6.3e-37;  
Matches 105; Conservative 43; Mismatches 109; Indels 17; Gaps 6  
QY 1 KPSSPPPELKFQCGK-----TLRPRKTIIGSGFTTITENQFWPAAYVRRHGGSVTVVCG 55  
DB 385 EPASFGQ---ACGRHKKTFLRPR--IIGSSSLPGSHPLAAIY---IGDS---FCA 433  
QY 56 GSLISPCWVISAHCIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLHKDYSADTL 115  
DB 434 GSLVHTCWVSAACHCFSHSPRDSVSVVLGQHFFNRTTDTQTFGIEKIPYTLVSFNP 493  
QY 116 AHNDIALLKIRSKEGCAQPSRTIQTICLPSVMYNDPQFGTSCBITGKENSVDLYPE 175  
DB 494 SDH-DLVLRUKKKGDRCATRSQVQICLPEPGSTFPAGHKQIAGWGLDENVSYS 552

```

176 QY QLKMTVVKVLSHRECOQPHYGVSEVYTKMLCAADPQWKTKSCQDSDSGPLVCSLQGRWTL 233
553 DB SLFEALVPLVADHKCSSPEVYGVADISPNWLKAGYFDCKSDACQDSDSGPLACRKGWYVL 612
236 QY TGIVSGRGCAKDKDPGVVTVRYSHLFWIRSHTK 269
613 DB YGIISWGDGGRHLKPGVTVRYVYVDWINDRIR 646

RESULT 19
ID12_CAVPO
FA12_CAVPO
ID ID_CAVPO STANDARD; PRT; 603 AA.
AC Q04962;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 13-MAR-2004 (Rel. 43, Last annotation update)
DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
DE (HAF) (Fragment).
DE F12.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
ON NCBI_TaxID=10141;
RX NCBI
CC -!- FUNCTION: Factor XII is a serum glycoprotein that participates in
CC the initiation of blood coagulation, fibrinolysis, and the
CC generation of bradykinin and angiotensin.
CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Ile bonds in factor
CC VII to form factor VIIa and factor XI to form factor Xla.
CC -!- MISCELLANEOUS: Factor XII, prekallikrein, and HMW kininogen form a
CC complex bound to an anionic surface. Prekallikrein is cleaved by
CC factor XII to form kallikrein, which then cleaves factor XII first
CC to alpha-factor Xla and then to beta-factor Xlaa. Alpha-factor
CC Xla activates factor XI to factor Xlaa.
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X68615; CAA48600.1; -.
CC FIR; S28941; S28941.
CC HSP; P00763; 1DPO.
CC
CC MEROPS: S01.211; -.
CC
CC InterPro; IPR003003; Cys_Ser_trypsin.
CC InterPro; IPR000742; EGF_2.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR006209; EGF_Like.
CC InterPro; IPR000083; Fibrinctnl.
CC InterPro; IPR000562; FN_Type_II.
CC InterPro; IPR006210; tEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Peptidase_S1.
CC InterPro; IPR001314; Peptidase_S1A.
CC Pfam; PF00008; EGF; 2.
CC Pfam; PF00039; fn1; 1.
CC Pfam; PF00040; fn2; 1.

```



QY 122 ALLKI-RSKEGRCAQPSRTIOTICLPMSYNDPQFG--TSCEITGFGKENSTDYLYPEQLK 178  
 Dp 449 ALLALQKADGSCAQLSPYVTCVCLPSGAPPSSSEFTTCCEVAGMGHGFEGAEYSSFLQ 508  
 QY 179 MTUVKLSHRECOQPHYVGVSTVTKMLCAADPQKWTDSCDGGGGLVC---SLQGRWTL 235  
 Dp 509 EAQPLISSEKSCSPFVHGDAFLGMLCAGFLGEGTDACQDGGGGLVCDEAAEHRLL 569  
 QY 236 TGIVSWGRCALKDKPGYVTVRSHFLPWRSHST 268  
 Dp 569 RGIVSWGSGCGDRNKPQVYTDVAGYLTWIKHT 601

RESULT 20  
 ID FAL2\_BOVIN STANDARD; PRT; 593 AA.  
 AC P98140;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)  
 DE (HAF) (Fragment).  
 GN F12.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver.  
 RX MEDLINE=94242782; PubMed=8186251;  
 RA Shibuya Y., Semba U., Okabe H., Kambara T., Yamamoto T.;  
 RT "Primary structure of bovine Hageman factor (blood coagulation factor  
 RT XII): comparison with human and guinea pig molecules.";  
 RL Biochim. Biophys. Acta 1206:63-70(1994).  
 RN [2]  
 RP SEQUENCE OF 10-21: 350-364 AND 525-550.  
 RX MEDLINE=77182112; PubMed=861210;  
 RA Fujikawa K., Walsh A.K., Davie W.E.;  
 RT "Isolation and characterization of bovine factor XII (Hageman  
 RT factor).";  
 RL Biochemistry 16:2270-2278(1977).  
 CC -!- FUNCTION: Factor XII is a serum glycoprotein that participates in  
 CC the initiation of blood coagulation, fibrinolysis, and the  
 CC generation of bradykinin and angiotensin.  
 CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Ile bonds in factor  
 CC VII to form factor VIIIa and factor XI to form factor XIa.  
 CC -!- PTM: O- AND N-GLYCOSYLATED (BY SIMILARITY).  
 CC -!- MISCELLANEOUS: Factor XII, prekallikrein, and HMW kininogen form a  
 CC complex bound to an anionic surface. Prekallikrein is cleaved by  
 CC factor XII to form kallikrein, which then cleaves factor XII first  
 CC to alpha-factor XIIa and then to beta-factor XIIa. Alpha-factor  
 CC XIIa activates factor XI to factor XIa. Bovine factor XII is  
 CC cleaved only to alpha-factor XIIa as it lacks the trypsin/  
 CC kallikrein cleavage site.  
 CC -!- SIMILARITY: Belongs to peptidase family S1.  
 CC -!- SIMILARITY: Contains 2 EGF-like domains.  
 CC -!- SIMILARITY: Contains 1 fibronectin type I domain.  
 CC -!- SIMILARITY: Contains 1 fibronectin type II domain.  
 CC -!- SIMILARITY: Contains 1 kringle domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; S70164; AAB30804.2; -;  
 CC PIR; S45281; S45281.  
 DR HSSP; P00763; LDPO.

DR MEROPS; S01.211; -;  
 DR InterPro; IPR009003; Cys Ser trypsin.  
 DR InterPro; IPR006209; EGF-like  
 DR InterPro; IPR000083; Fibinectin  
 DR InterPro; IPR000562; FN\_Type\_II.  
 DR InterPro; IPR006210; IEGF.  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR Pfam; PF00008; EGF; 2.  
 DR Pfam; PF00039; fn1; 1.  
 DR Pfam; PF00040; fn2; 1.  
 DR Pfam; PF00051; kringle; 1.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PRO0722; CHYMOTRYPSIN.  
 DR PRINTS; PRO0013; FNTYPEII.  
 DR PRINTS; PRO0018; KRINGLE.  
 DR ProDom; PD000995; FN\_Type\_II; 1.  
 DR ProDom; PD000395; Kringle; 1.  
 DR SMART; SM00181; EGF; 2.  
 DR SMART; SM00058; FN1; 1.  
 DR SMART; SM00059; FN2; 1.  
 DR SMART; SM00130; KR; 1.  
 DR SMART; SM00020; Tryp\_Spc; 1.  
 DR PROSITE; PS00022; EGF\_1; 2.  
 DR PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
 DR PROSITE; PS00026; EGF\_3; 2.  
 DR PROSITE; PS01253; FIBRONECTIN\_1; 1.  
 DR PROSITE; PS00023; FIBRONECTIN\_2; 1.  
 DR PROSITE; PS00021; KRINGLE\_1; 1.  
 DR PROSITE; PS00070; KRINGLE\_2; 1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;  
 KW Hydrolyase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.  
 FT NON\_TER 1  
 FT SIGNAL <1 9  
 FT CHAIN 10 349  
 FT CHAIN 350 593  
 FT DOMAIN 37 78  
 FT DOMAIN 84 121  
 FT DOMAIN 123 163  
 FT DOMAIN 164 200  
 FT DOMAIN 207 287  
 FT DOMAIN 297 333  
 FT DOMAIN 350 593  
 FT ACT\_SITE 389 389  
 FT ACT\_SITE 438 438  
 FT ACT\_SITE 541 541  
 FT DISULFID 88 100  
 FT DISULFID 94 109  
 FT DISULFID 111 120  
 FT DISULFID 125 153  
 FT DISULFID 151 160  
 FT DISULFID 168 179  
 FT DISULFID 173 188  
 FT DISULFID 190 199  
 FT DISULFID 207 287  
 FT DISULFID 230 269  
 FT DISULFID 258 282  
 FT DISULFID 336 463  
 FT DISULFID 374 390  
 FT DISULFID 382 452  
 FT DISULFID 413 416  
 FT DISULFID 479 547  
 FT DISULFID 510 526  
 FT DISULFID 537 568  
 FT CARBOHYD 99 99  
 FT CARBOHYD 241 241  
 FT CARBOHYD 263 263  
 FT CARBOHYD 410 410  
 FT SEQUENCE 593 AA; 65148 MW; 721952BA792ED61F CRC64;



Query Match 30.0%; Score 453; DB 1; Length 593;  
Best Local Similarity 38.5%; Pred. No. 2.6e-33;  
Matches 102; Conservative 44; Mismatches 101; Indels 18; Gaps 7;

QY 13 CQ---KTLRPFKIGGFTTIENQWFAALYRHRGGSVTVYCGSLIPFCWVISATH 69  
DB 336 CQRLRKWLSNLVVGVALPGHPIYALYDQ-----HFCAGSLIAPCWVLTAAH 389

QY 70 CFIDVPKEDYIVYGRNLNNTGEMKFEVENILHKDYSADTLAHNDIALKIR-S 128  
DB 390 CLQNRPAKELVWLQGRHNSCEQCOTLAVRDYRLHFAFSPITYQH--DLALVRLQES 447

QY 129 KEGRCAPQRTTQICLPSVYNDPOFGNS--CEITGFGKNSDVLYPEQLKMTVWKLIS 186  
DB 448 ADGCCAPSPFPVQVCLDSTAAAPAESEAAVCEVAGHQPGEGB-YSFLOEAQVPLID 506

QY 187 HRECOQPHYGVSEVTTKMLCAADPOWKTDSCGDSGGPLVC---SLQGRMTLTGIVSWGR 243  
DB 507 PQRCAPDVHGAATQGMKACGFLGCTDACQDGGPLVCBDETPERQLILRGIVSNGS 566

QY 244 GCALKDKPGVTVSHFLPWIRSH 268  
DB 567 GCGNRLKPGVYTDVANYLAWIREHT 591

RESULT 21  
NETR HUMAN STANDARD; PRT; 875 AA.

ID NETR HUMAN STANDARD; PRT; 875 AA.  
AC P56730; Q9UP16;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Neurotysin precursor (BC 3.4.21.-) (Motopsin) (Leydin).  
GN PRSS12.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP TISSUE=Brain;  
RC MEDLINE=98201705; PubMed=9540828;  
RX Proba K., Gschwend T.P., Sonderegger P.;  
RA "Cloning and sequencing of the cDNA encoding human neurotysin.";  
RL Biochim. Biophys. Acta 1396:143-147(1998).  
RN [2]  
RP SEQUENCE OF 615-875 FROM N.A.  
RC TISSUE=Testis;  
RX MEDLINE=99203523; PubMed=10103056;  
RA Poorafshar M., Hellman L.;  
RT "Cloning and structural analysis of leydin, a novel human serine  
protease expressed by the Leydig cells of the testis.";  
RL Eur. J. Biochem. 261:244-250(1999).  
CC -1- FUNCTION: Plays a role in neuronal plasticity and the proteolytic  
action may subserve structural reorganizations associated with  
learning and memory operations (By similarity).  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: Brain and Leydig cells of the testis.  
CC -1- SIMILARITY: Belongs to peptidase family S1.  
CC -1- SIMILARITY: Contains 1 kringle domain.  
CC -1- SIMILARITY: Contains 4 SRCR domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AJ001531; CAA04816.1; --  
DR EMBL; AF07298; AAD25919.1; --

HSSP; P00763; LDPO.  
Genew; HGNC:9477; PRSS12.  
DR MIM; 606709; --  
DR MEROPS; S01.237; --  
DR GO; GO:008236; F:serine-type peptidase activity; TAS.  
DR InterPro; IPR009003; Cys Ser trypsin.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR001254; Peptidase S1.  
DR InterPro; IPR001314; Peptidase SLA.  
DR InterPro; IPR001190; Srcr\_receptor.  
DR Pfam; PF00051; Kringle; 1.  
DR Pfam; PF00530; SRCR; 4.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00258; SPERACTRCPTR.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00202; SR; 4.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS00070; KRINGLE\_2; 1.  
DR PROSITE; PS00420; SRCR\_1; 3.  
DR PROSITE; PS00287; SRCR\_2; 4.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease; Glycoprotein; Kringle; Repeat; Signal.  
FT SIGNAL 1 20  
FT CHAIN 21 875  
FT DOMAIN 23 92  
FT DOMAIN 93 165  
FT DOMAIN 170 271  
FT DOMAIN 280 381  
FT DOMAIN 387 487  
FT DOMAIN 500 601  
FT DOMAIN 619 875  
FT DOMAIN 619 630  
FT ACT\_SITE 630 631  
FT ACT\_SITE 676 676  
FT ACT\_SITE 726 726  
FT ACT\_SITE 825 825  
FT DISULFID 619 750  
FT CARBOHYD 26 26  
FT CARBOHYD 683 683  
FT CONFLICT 563 563  
FT CONFLICT 701 701  
FT CONFLICT 839 841  
SQ SEQUENCE 875 AA; 97011 MW; B66EC946DC208DCB CRC64;

Query Match 29.3%; Score 442.5; DB 1; Length 875;  
Best Local Similarity 36.6%; Pred. No. 3.8e-32;  
Matches 100; Conservative 47; Mismatches 113; Indels 13; Gaps 8

QY 3 SSPPEELKFCQGTLPKPK-IIGGFTTIENQWFAALYRHRGGSVTVYCGSLIP 61  
DB 609 NSNKESLSVCGLLHRRQKRIIGKNSLRGGWPMQVSLRSLKSHGDSGLCCATLLSS 668

QY 62 CWVLSATHCFIDYPKK-EDIVYLGSRSLNSNTGEMKFEVENILHKDYSADTLAHND 120  
DB 669 CWVLSATHCFIDYPKK-EDIVYLGSRSLNSNTGEMKFEVENILHKDYSADTLAHND 120

QY 121 IALLKIRSKGRCAQPSRTTQICLPSVYNDPO--FGTSCITGFGKNSDVLYPEQLK 179  
DB 727 IALLKIRSKGRCAQPSRTTQICLPSVYNDPO--FGTSCITGFGKNSDVLYPEQLK 179

QY 180 TVVKLISHRECOQPHYGVSEVTTKMLCAAD--POWKTDSCGDSGGPLVCISLQGR--MTLT 236  
DB 784 AAIFLLPKRFCEE--RYNGRPTGRMLCAGNLHFKRVDSQCGDSGGPLMCRPQESVNVY 841

QY 237 GIVSWGRCAKDKPGVTVSHFLPWIRSH 269  
DB 842 GIVSWGRCAKDKPGVTVSHFLPWIRSH 269

```
RESULT 22
NETR MOUSE
ID NETR MOUSE STANDARD; PRT; 761 AA.
AC O08762;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Neurotysin precursor (EC 3.4.21.-) (Motopsin) (Brain-specific serine
GN PRS12 OR BSSP3).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
SEQUENCE FROM N.A.
RP TISSUE=Brain;
RX MEDLINE=97401523; PubMed=9245503;
RA Gschwend T.P., Krueger S.R., Kozlov S.V., Wolfer D.P., Sonderegger P.,
RT "Neurotysin, a novel multidomain serine protease expressed in the
RT nervous system.",
RL Mol. Cell. Neurosci. 9:207-219(1997).
RN [2]
SEQUENCE FROM N.A.
RP MEDLINE=98008948; PubMed=9344839;
RA Yamamura Y., Yamashiro K., Tsuruoka N., Nakazato H., Tsujimura A.,
RT "Molecular cloning of a novel brain-specific serine protease with a
RT kringle-like structure and three scavenger receptor cysteine-rich
RT motifs.",
RL Biochem. Biophys. Res. Commun. 239:386-392(1997).
RN [3]
SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Mammary gland;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Moquellano N.A., Peters G.J., Carninci P., Prange C.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,
RA Schnerch A., Schein J.E., Jones S.U.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.",
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Plays a role in neuronal plasticity and the proteolytic
CC action may subserve structural reorganizations associated with
CC learning and memory operations.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Most abundant in cerebral cortex, hippocampus
CC and amygdala.
CC -1- SIMILARITY: Belongs to peptidase family S1.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -1- SIMILARITY: Contains 3 SRCR domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC
DR EMBL; Y13192; CAAT3646.1; -.
DR EMBL; D89871; BAA23986.1; -.
DR EMBL; BC031429; AAH31429.1; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.237; -.
DR MGD; MGI:1100881; Prss12.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kringle_1.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00051; Kringle_1.
DR Pfam; PF00530; SRCR_3.
DR Pfam; PF00089; trypsin_1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00258; SPERACINCPTR.
DR ProDom; PD000395; Kringle_1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00202; SR; 3.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00420; SRCR_1; 3.
DR PROSITE; PS00287; SRCR_2; 3.
DR PROSITE; PS00240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hydrolase; Serine protease; Glycoprotein; Kringle; Repeat; Signal.
FT SIGNAL; 1 21 POTENTIAL.
FT CHAIN; 22 761 NEUOTRYPsin.
FT DOMAIN; 85 157 KRINGLE.
FT DOMAIN; 166 267 SRCR 1.
FT DOMAIN; 273 373 SRCR 2.
FT DOMAIN; 386 487 SRCR 3.
FT DOMAIN; 505 761 SERINE PROTEASE.
FT DOMAIN; 505 516 ZYMOGEN ACTIVATION REGION.
FT ACT_SITE; 516 517 REACTIVE BOND (POTENTIAL).
FT ACT_SITE; 562 562 CHARGE RELAY SYSTEM.
FT ACT_SITE; 612 612 CHARGE RELAY SYSTEM.
FT ACT_SITE; 711 711 CHARGE RELAY SYSTEM.
FT DISULFID; 505 636 POTENTIAL.
FT CARBOHYD; 93 93 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD; 521 521 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD; 569 569 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 761 AA; 84118 MW; DF507B03712164E6 CRC64;

Query Match 28.2%; Score 425.5; DB 1; Length 761;
Best Local Similarity 37.1%; Pred. No. 1.1e-30;
Matches 101; Conservative 42; Mismatches 116; Indels 13; Gaps 8

Qy 3 SSPPEELKFCGOKTLRPRFK-IIGGEFTTIENQWPAAYRRHRGGSVTVCGSLISP 61
Db 495 SGKEMLSGCCGLRLHRRQKRIIGNNSLRGAWPQASLRLESAGHGDGLLCGATLSS 554
Qy 62 CWVISATPCFIDY-PKEDYIVYVYLGSRSLNSNTQGMKPEVENLILKQYSDTLAHND 120
Db 555 CWLTAHACFKRYGNNRSYAVRVGDVHTLVPEEFQEIQVQVIVHRYRPRDSY--D 612
Qy 121 IALLKTESGRCAQPSRTIQTICLPSMYNDPO-FGTSCEITGFGKENSTDYLYPEOLKM 179
Db 613 IALVRLGQGEQCARLSTHVLPACLPLWRERPKTASNCIHWG---DTGRAYSRITLQ 669
Qy 180 TVVKLIISHRECQQPHYYGSEVTTKMLCAADPQW--KTDSCQGSQGGPLVCSLQR-MTLT 236
Db 670 AAVELLPRFCKE--RYKGLFTGRMLCAGNLQEDNRVDSQCGSGGSLMCKEKPDESVMVY 727
Qy 237 GIVSWGGCCALKDQPGVYTVSVHPLWIRSH 268
Db 728 GVTSGWGYCGGVKDPFGVYTVRVPFVPMIKSVT 759

RESULT 23
```

HATT\_HUMAN  
 ID HATT\_HUMAN STANDARD; PRT; 418 AA.  
 AC 060235;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Airway trypsin-like protease precursor (EC 3.4.21.-).  
 GN HAT.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. PubMed=9565616;  
 RX MEDLINE=98234382; PubMed=9565616;  
 RA Yasuoka K., Masuda K.-I., Ogawa H., Takagi K., Umemoto N., Yasuoka S.;  
 RT "Cloning and characterization of the cDNA for human airway trypsin-  
 like protease.";  
 RL J. Biol. Chem. 273:11895-11901(1998).  
 RN [2]  
 RP SEQUENCE OF 187-206, AND CHARACTERIZATION.  
 RX MEDLINE=97224034; PubMed=9070615;  
 RA Masuda K., Onishi T., Kawano S., Teuchihashi S., Ogawara M.,  
 RT "Purification, characterization, and localization of a novel  
 trypsin-like protease found in the human airway.";  
 RL Am. J. Respir. Cell Mol. Biol. 16:300-308(1997).  
 CC -!- FUNCTION: May play some biological role in the host defense system  
 on the mucous membrane independently of or in cooperation with  
 other substances in airway mucous or bronchial secretions.  
 CC -!- CATALYTIC ACTIVITY: Preferentially cleaves the C-terminal side of  
 arginine residues at the P1 position of certain peptides, cleaving  
 Boc-Phe-Ser-Arg-4-methylcoumaryl-7-amide most efficiently and  
 having an optimum pH of 8.6 with this substrate.  
 CC -!- ENZYME REGULATION: Strongly inhibited by diisopropyl  
 fluorophosphate, leupeptin, antipain, aprotinin, and soybean  
 trypsin inhibitor, but hardly inhibited by secretory leukocyte  
 protease inhibitor at 10 microM.  
 CC -!- SUBUNIT: Monomer.  
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Activated by  
 cleavage and secreted.  
 CC -!- TISSUE SPECIFICITY: Located in the cells of the submucosal serous  
 glands of the bronchi and trachea.  
 CC -!- SIMILARITY: Belongs to peptidase family S1.  
 CC -!- SIMILARITY: Belongs to SEA domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AB002134; BAA28691.1; --  
 DR HSP; P00750; 1RTF.  
 DR MEROPS; S01.301; --  
 DR MIM; 605369; --  
 DR GO; GO:0005576; C:extracellular; TAS.  
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.  
 DR GO; GO:0008233; F:peptidase activity; TAS.  
 DR GO; GO:0007585; P:respiratory gaseous exchange; TAS.  
 DR InterPro; IPR009003; Cys Ser trypsin.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR000082; SEA domain.  
 DR Pfam; PF01390; SEA; 1.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PRO0722; CHYMOTRYPSIN.  
 DR SMART; SM00200; SEA; 1.  
 DR SMART; SM00200; Tryp\_Spc; 1.  
 DR PROSITE; PSS0024; SEA; 1.  
 DR PROSITE; PSS0240; TRYPsin\_DOM; 1.

DR PROSITE; PSS00134; TRYPsin\_HIS; 1.  
 DR PROSITE; PSS00135; TRYPsin\_SER; 1.  
 KW Hydrolase; Serine protease; Transmembrane; Signal-anchor; Zymogen;  
 FT CHAIN 1 186 AIRWAY TRYPsin-LIKE PROTEASE, NON-  
 CATALYTIC CHAIN.  
 FT CHAIN 187 418 AIRWAY TRYPsin-LIKE PROTEASE, CATALYTIC  
 CHAIN.  
 FT DOMAIN 1 20 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 21 41 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 (POTENTIAL).  
 FT DOMAIN 42 418 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 44 164 SEA.  
 FT DOMAIN 187 417 SERINE PROTEASE.  
 FT ACT\_SITE 227 272 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 272 272 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 368 368 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT DISULFID 173 292 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 212 228 BY SIMILARITY.  
 FT DISULFID 337 353 BY SIMILARITY.  
 FT DISULFID 364 393 BY SIMILARITY.  
 FT CARBOHYD 144 144 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 418 AA; 46263 MW; F4BC1DB020CFBBD0 CRC64;  
 Query Match 28.1%; Score 423.5; DB 1; Length 418;  
 Best Local Similarity 36.4%; Pred. No. 7.9e-31;  
 Matches 90; Conservative 47; Mismatches 93; Indels 17; Gaps 6;  
 QY 23 KIIGSEFTIENQFWAAIYRRHGGSVTVVCGSLISPCWVISAHCIDYPKEDYIV 82  
 DB 186 RILGGTEAEEGSPWQVSLANN-----AHHCGLINNNWILTAACFRSNPNRDWIA 240  
 QY 83 YLCGRSLNNTQEMKFEVENLTHQVSDATLHNDIALLKIRKGCRCQAQPSRTIQT 142  
 DB 241 TSG-----ISTTFPKLRVRNLIHNNYKAT--HENDIALVRLNS----VTFTKDIHS 290  
 QY 143 ICLPSMNDPQFQTSCEITGFKENSTDYLPQLKMTVVVKLIHRECOQPHYGVSEVTT 202  
 DB 291 VCLPAATQNPSTAYTVTGWGAQYAGHTVPE-LRQGVRIISNDVNCNAPHSYNGAII 349  
 QY 203 KMLCAADPQWTKDCGDSGLVCSLQGRM-TLTGIVSGRGCAKDKPGVYVTRVSHFL 261  
 DB 350 GMLCAGVPQGVGDACGDSGLVQSDRLWLFVIGVSWGQCGLPDRPGVYVTRVAYL 409  
 QY 262 FWIRSH 268  
 DB 410 DWIRQQT 416  
 RESULT 24  
 PSS8\_HUMAN STANDARD; PRT; 343 AA.  
 ID PSS8\_HUMAN  
 AC Q16651; Q9UCA3;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Prostatin precursor (EC 3.4.21.-).  
 GN PSS8.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Prostate;  
 RX MEDLINE=95286644; PubMed=7768952;  
 RA Yu J.X., Chao L., Chao J.;  
 RT "Molecular cloning, tissue-specific expression, and cellular  
 localization of human prostatin mRNA.";  
 RL J. Biol. Chem. 270:13483-13489(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;

RX MEDLINE=22389257; PubMed=12477932;  
 RA Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullane S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalus D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [3]  
 RP SEQUENCE OF 45-64.  
 RC TISSUE=Semen;  
 RX MEDLINE=94308140; PubMed=8034638;  
 RA Yu J.X., Chao L., Chao J.;  
 RT "Prostatein is a novel human serine proteinase from seminal fluid.  
 RT Purification, tissue distribution, and localization in prostate  
 gland";  
 RL J. Biol. Chem. 269:18843-18848(1994).  
 CC -1- FUNCTION: Possesses a trypsin-like cleavage specificity.  
 CC -1- SUBUNIT: Heterodimer of two chains, light and heavy, held by a  
 disulfide bond.  
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND. SECRETED AFTER CLEAVAGE OF  
 ITS C-TERMINUS.  
 CC -1- TISSUE SPECIFICITY: Found in prostate, liver, salivary gland,  
 kidney, lung, pancreas, colon, bronchus and renal proximal tubular  
 cells. In the prostate gland it may be synthesized in epithelial  
 cells, secreted into the ducts, and excreted into the seminal  
 fluid.  
 CC -1- SIMILARITY: Belongs to peptidase family S1.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; L41351; AAC41759.1; -;  
 DR EMBL; U33446; AAB19071.1; -;  
 DR EMBL; BC001462; AAB19071.1; -;  
 DR PIR; A57014; A57014.  
 DR HSP; P00763; IDPO.  
 DR MEROPS; S01.159; -;  
 DR Gene; HGNC:9491; PRSS8.  
 DR MAM; 600823; -;  
 DR GO; GO:0005615; C:extracellular space; TAS.  
 DR GO; GO:0005886; C:plasma membrane; TAS.  
 DR GO; GO:0008236; F:serine-type peptidase activity; TAS.  
 DR InterPro; IPR009003; Cys Ser trypsin.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; TRYP\_SPC; 1.  
 DR PROSITE; PS0240; TRYP\_SIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYP\_SIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYP\_SIN\_SER; 1.  
 KW Hydrolase; Serine protease; Zymogen; Signal; Glycoprotein;  
 KW Transmembrane.  
 FT SIGNAL 1 29 POTENTIAL.

FT PROPEP 30 32 ACTIVATION PEPTIDE.  
 FT CHAIN 33 44 PROSTATIN LIGHT CHAIN.  
 FT CHAIN 45 322 PROSTATIN HEAVY CHAIN.  
 FT PROPEP 323 343 POTENTIAL.  
 FT TRANSMEM 320 340 SERINE PROTEASE.  
 FT DOMAIN 45 286 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 37 154 BY SIMILARITY.  
 FT DISULFID 70 86 BY SIMILARITY.  
 FT DISULFID 168 244 BY SIMILARITY.  
 FT DISULFID 201 223 BY SIMILARITY.  
 FT DISULFID 234 262 BY SIMILARITY.  
 FT ACT\_SITE 85 85 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 134 134 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 238 238 CHARGE RELAY SYSTEM.  
 FT CARBOHYD 159 159 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 343 AA; 98DD6447F5A8C1B2 CRC64;  
 Query Match 27.1%; Score 408.5; DB 1; Length 343;  
 Best Local Similarity 36.6%; Pred. No. 1.4e-29;  
 Matches 100; Conservative 36; Mismatches 100; Indels 37; Gaps 8

QY 13 CGQKTLRPKIIIGGFTTIENQNPFAALYRRHGGSVTY----VCGGSLIPCWVISAT 68  
 DB 37 CG---VAPOARITGSSAVAGQWPQV-----SITYEGVHVGSSLVSEQWYLSAA 84  
 QY 69 HCFIDYPKKEDYIVYLGSRGLNSNTQEMKFEVENLILHKDYSDATLAHNDIALKIRS 128  
 DB 85 HCFPSEHKEAYEVKLGAGHQLDSYSDAKVSLKDIIPHSYLOE--GSGDIALQL-- 140  
 QY 129 KEGRCAQPSRTIOTICLPSMYNDPQGTSCETIGFK-ENSTDYLYPEOLKMTVVKLISH 187  
 DB 141 --SRPTFSYIRPICLPAANASFPNGLHCTVTGHWGVAPSVSLTTPKLPQCLEVFLISR 196  
 QY 188 REC-----QPHYGVSEVTTKMLCAADPQWKTDCQGDGSGGFLVCSLQGRWTLTG 237  
 DB 199 ETNCNLYINDAKPEEPHF-----VOEDMVCAGYEGKDACQDGSGLSPVGLWYLTG 254  
 QY 238 IVSGRGCAKDKPGVYTRVSHFLPWRSHKE 270  
 DB 255 IVSGWDACGARNRPGVYTLASSYASWTQSKYTE 287

RESULT 25  
 KAL\_MOUSE STANDARD; PRT; 638 AA.  
 ID KAL\_MOUSE  
 AC P26262;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41; Last annotation update)  
 DE Plasma kallikrein precursor (EC 3.4.21.34) (Plasma prekallikrein)  
 DE (Kininogenin) (Fletcher factor).  
 GN KLB1 OR KLB3 OR PK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC STRAIN=BALB/c; TISSUE=Liver;  
 RX MEDLINE=9109044; PubMed=2264928;  
 RA Seidah N.G., Sawyer N., Hamelin J., Mion P., Beaubien G.,  
 RA Brachpala L., Rochemont J., Molikay M., Chretien M.;  
 RT "Mouse plasma kallikrein: cDNA structure, enzyme characterization,  
 RT and comparison of protein and mRNA levels among species";  
 RL DNA Cell Biol. 9:737-748(1990).  
 CC -1- FUNCTION: The enzyme cleaves Lys-Arg and Arg-Ser bonds. It  
 activates, in a reciprocal reaction, factor XII after its binding  
 to a negatively charged surface. It also releases bradykinin from  
 HMW kininogen and may also play a role in the renin-angiotensin  
 system by converting protein into renin.  
 CC -1- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Xaa and Lys-|-Xaa  
 bonds, including Lys-|-Arg and Arg-|-Ser bonds in (human)  
 kininogen to release bradykinin.



RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
RA Shiraki T., Waki K., Kawai J., Aizawa K., Akakawa T., Fukuda S.,  
RA Hara A., Hashizume W., Motani K., Ishii Y., Itoh M., Kagawa I.,  
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
RA Birney E., Hayashizaki Y.,  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs";  
RL Nature 420:563-573(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klusner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zellers B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,  
RA Brownstein M.J., Ustin T.B., Itoh Yuki S., Carninci P., Frange C.,  
RA Raha S.S., Loquellano N.A., Peters G.U., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Rahay J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.B.,  
RA Schnerch A., Schein J.E., Jones S.J., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
RL and mouse cDNA sequences";  
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RP [4]  
RX MEDLINE=22668120; PubMed=12784999;  
RA Netzel-Arnett S., Hooper J.D., Szabo R., Madison E.L., Quigley J.P.,  
RA Bugge T.H., Antalio T.M.,  
RT "Membrane anchored serine proteases: a rapidly expanding group of cell  
RT surface proteolytic enzymes with potential roles in cancer";  
RL Cancer Metastasis Rev. 22:237-258(2003).  
CC -!- FUNCTION: May play a specialized role in matrix remodeling  
CC processes in liver (By similarity).  
CC -!- SUBCELLULAR LOCATION: Type II membrane protein.  
CC -!- TISSUE SPECIFICITY: Expressed at highest levels in liver, kidney  
CC and uterus.  
CC -!- DEVELOPMENTAL STAGE: Expressed at higher levels from 12.5 dpc to  
CC 15.5 dpc with a peak at 13.5 dpc. Expression in the developing  
CC liver as well as a restricted set of embryonic epithelial cells of  
CC the nasal cavity and pharyngo-tympenic tubes.  
CC -!- SIMILARITY: Belongs to peptidase family 11.  
CC -!- SIMILARITY: Contains 2 CUB domains.  
CC -!- SIMILARITY: Contains 3 LDL-receptor class A domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AY240929; AAF69827.1; -  
CC EMBL; AK004939; BAB23884.2; -  
CC EMBL; BC029645; AAB29645.2; -  
CC HSSP; P00763; 1DPO.  
CC MEROPS; S01.308; -  
CC MGD; MGI:1919003; TMRPSS6.  
CC InterPro; IPR000859; CUB.  
CC InterPro; IPR009003; Cys\_Ser\_trypsin.  
CC InterPro; IPR00172; LDL\_receptor\_A.  
CC InterPro; IPR001254; Peptidase\_S1.

DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00057; ldl\_recept\_a; 2.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00261; LDLRECEPTOR.  
DR PROSITE; PS01180; CUB\_1.  
DR PROSITE; PS01180; CUB\_2.  
DR PROSITE; PS01209; LDLRA\_1.  
DR PROSITE; PS01209; LDLRA\_2.  
DR PROSITE; PS01209; LDLRA\_3.  
DR PROSITE; PS02040; TRYPSIN\_DOM; 1.  
DR PROSITE; PS02040; TRYPSIN\_HIS; 1.  
DR PROSITE; PS01134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS01135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease; Repeat; Signal-anchor; Transmembrane;  
KW Glycoprotein.  
FT DOMAIN 1 59 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 60 80 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
(POTENTIAL).  
FT DOMAIN 81 811 EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 213 336 CUB 1.  
FT DOMAIN 323 440 CUB 2.  
FT DOMAIN 445 477 LDL-RECEPTOR CLASS A 1.  
FT DOMAIN 478 514 LDL-RECEPTOR CLASS A 2.  
FT DOMAIN 518 555 LDL-RECEPTOR CLASS A 3.  
FT DOMAIN 565 799 SERINE PROTEASE.  
FT ACT\_SITE 617 671 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 668 668 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 762 762 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT CARBOHYD 138 138 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 216 216 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 338 338 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 433 433 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 453 453 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 518 518 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CONFLICT 690 690 P -> PP (IN REF. 2).  
SQ SEQUENCE 811 AA; 90978 MW; 32EB3E7C127801B CRC64;  
  
Query Match 27.0%; Score 406.5; DB 1; Length 811;  
Best Local Similarity 36.7%; Pred. No. 6.2e-29;  
Matches 99; Conservative 47; Mismatches 89; Indels 35; Gaps 10.  
  
QY 7 BELKFGQGTLPKPRKIIIGCEPTTIENQWPAIYRRHGGSVTVVCGSLSPCWVIS 66  
DB 562 DEQCHDCGLQGLSSR--IVGTVSSEGEWPMQASLQIRG-----HICGALLADRAWIT 614  
  
QY 67 ATHCFIDYPKED-----YIVLGRSLNNTQGMKPEVENLILHKDYSADTLAHN 119  
DB 615 AAHCF-----QEDSMASPKLWTVFLGKMRQNSRWPGEVSFKVSRLLFHPVHEED--SHDY 667  
  
QY 120 DIALLKIRSKGRCAQP---SRTQITCLPSMYNDPQGTSCBITGKNSDYLYPEQ 176  
DB 668 DVALQL-----DHPVTVYSAIVRVCVLPARSHFFEPGQHCWITGWAQREGGPV-SNT 719  
  
QY 177 LKMTVVKLISHRECCQPHYYGSEVTKLCAADPQWTKDCQGDGSGPLVC-SLQGRWTL 235  
DB 720 LQKVDVQLVPODLGSEAYR--QVSPRLCAGYRKGGKDKACQGDGSGPLVCPREPSGRWFL 777  
  
QY 236 TGIYSWGRGCKADKPGVTVRVSHFLPWIR 265  
DB 778 AGLVSWGLGCGRPNFPGVYTRVTRVINWIQ 807  
  
RESULT 27  
KLKD HUMAN  
ID KLKD HUMAN STANDARD; PRT; 277 AA.  
AC Q9UKR3; Q9Y433;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Kallikrein 13 precursor (EC 3.4.21.-) (Kallikrein-like protein 4)  
DE (KLK-L4).  
GN KLK13 OR KLK14.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20229789; PubMed=10766816;  
 RA Yousef G.M., Chang A., Diamandis E.P.;  
 RT "Identification and characterization of KIK-14, a new kallikrein-like  
 RT gene that appears to be down-regulated in breast cancer tissues";  
 RL J. Biol. Chem. 275:11891-11898(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,  
 RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S.,  
 RA Phan H., Velasco N., Do L., Regala W., Terry A., Brower A., Garnes J.,  
 RA Pandangan L., Exler A., Christensen M., Georgescu A., Avila J., Liu S.,  
 RA Andreise T., Frankheim M., Attix C., Anico-Keller G., Coefield J.,  
 RA Duarte S., Lucas R., Bruce R., Thomas P., Quan G., Kronmiller B.,  
 RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,  
 RA Olsen A.S., Carrano A.V.;  
 RT "Sequence analysis of chromosome 19q13.4";  
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 1-180 FROM N.A.  
 RC TISSUE=Uterus;  
 RA Ansoerge W., Wirkner U., Mewes H.-W., Gassenhuber J., Wiemann S.;  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).  
 CC -1- TISSUE SPECIFICITY: Expressed in prostate, breast, testis and  
 CC salivary gland.  
 CC -1- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.  
 CC  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; AF135024; AAD26425.2; -;  
 DR EMBL; AC011473; AAG23259.1; -;  
 DR EMBL; AL050220; CAB43320.1; ALT\_INIT.  
 DR HSSP; P00763; IDPO.  
 DR MEROPS; S01.306; -;  
 DR Genew; HGNC:6361; KLIK3.  
 DR MIM; 605505; -;  
 DR GO; GO:0005576; C:extracellular; NAS.  
 DR GO; GO:0004252; F:serine-type endopeptidase activity; NAS.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; NAS.  
 DR InterPro; IPR009003; Cys Ser trypsin.  
 DR InterPro; IPR01254; Peptidase\_S1.  
 DR InterPro; IPR01314; Peptidase\_S1A.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; TRYP\_SPC; 1.  
 DR PROSITE; PS02040; TRYP\_SIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYP\_SIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYP\_SIN\_SER; 1.  
 KW Hydrolase; Serine protease; Glycoprotein; Signal.  
 FT SIGNAL 1 16  
 FT CHAIN 17 277  
 FT ACT\_SITE 76 76  
 FT ACT\_SITE 124 124  
 FT ACT\_SITE 218 218  
 FT ACT\_SITE 218 218  
 FT DISULFID 42 178  
 FT DISULFID 61 77  
 FT DISULFID 157 224  
 FT DISULFID 189 203  
 FT DISULFID 214 239  
 FT CARBOHYD 30 30  
 FT CARBOHYD 225 225  
 FT CONFLICT 170 180 VNPFTLQCAN --> GWHPHRPEAP (IN REF. 3).  
 SEQUENCE 277 AA; 30570 MW; BA8A9E8DCFB5D542 CRC64;

Query Match 26.9%; Score 405; DB 1; Length 277;  
 Best Local Similarity 41.1%; Pred. NO. 2.3e-29;  
 Matches 101; Conservative 34; Mismatches 81; Indels 30; Gaps 10  
 QY 26 GGEFTTIENOPFAAIVRRHGGSVTVVCGGSLSPQWISATHCFIDYPKKEDYIVYLG 85  
 DB 38 GGYTCFPHSPQWAAALVQGR-----LLCGGVVHPKWLTAHCL-----KEGLKYLG 87  
 QY 86 RSRNSNTQGMKFEVENLILHKDY--SADTLAHNDIALLKTRSKRGCAQSPRTIQT 143  
 DB 88 KHALGRVEAGEQVREVVHSIPHPHYRRSPHLNHDHDMLELQSP-----VOLTGYIQT- 142  
 QY 144 CLPSMYND---PQFGTSCITGKGNSTDYLPKQMTVVKLISHRECCQPHYTGSEV 200  
 DB 143 -LPUSHNRITP--GITCRVSGWGTTSPOVNTPKLQCANIQURDEECRQ--VTPGKI 197  
 QY 201 TTKMLCAADPQWKTDSCQSGGSLVCSLQGRMTLFGIVSWGR--GCALKDKPKGVYTRVSH 259  
 DB 198 TDNMLCAGTKEGGKDCSGSGGPLVCN----RTLYGIVSGWGFPGQPPRGVYTRVSR 253  
 QY 260 FLPWIR 265  
 DB 254 YVLWIR 259  
 RESULT 28  
 TMS6 HUMAN STANDARD; PRT; 811 AA.  
 ID TMS6 HUMAN STANDARD; PRT; 811 AA.  
 AC Q81U80; Q81U82; Q81XV8;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Transmembrane protease, serine 6 (EC 3.4.21.-) (Matriptase-2).  
 GN TMRPS6.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RA Hooper J.D., Quigley J.P.;  
 RT "TMRPS6, a new type II transmembrane serine protease";  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20057165; PubMed=10591208;  
 RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,  
 RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,  
 RA Baguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,  
 RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,  
 RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,  
 RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,  
 RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,  
 RA Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,  
 RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,  
 RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,  
 RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,  
 RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,  
 RA Laird G.K., Langford C.F., Levensha M.A., Lloyd C., Lloyd D.M.,  
 RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,  
 RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,  
 RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,  
 RA Phillips S.H., Plumb R.W., Ramsey H., Ramsey Y., Rogers L., Ross M.T.,  
 RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L., Swann R.M.,  
 RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,  
 RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,  
 RA Williams L., Williams S.A., Williamson H., Wilmer T.B., Wilming L.,  
 RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,  
 RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,  
 RA Shintani A., Shibusu K., Yoshizaki Y., Aoki N., Mitsuana S.,  
 RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,  
 RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,



This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce>, or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AJ319876; CAC85953.1; ALT INIT.  
EMBL; AY055383; AAL16413.1; -  
EMBL; AY055384; AAL16414.1; -  
EMBL; AL022314; -; NOT ANNOTATED\_CDS.  
EMBL; BC039082; AAH39082.1; -  
InterPro; IPR000859; CUB 1.  
InterPro; IPR009003; Cys\_Ser\_trypsin.  
InterPro; IPR002172; LDL\_receptor\_A.  
InterPro; IPR001254; Peptidase\_S1.  
InterPro; IPR001314; Peptidase\_S1A.  
Pfam; PF00431; CUB; 1.  
Pfam; PF00057; ldl\_recept\_a; 2.  
Pfam; PF00089; trypsin\_1\_  
PRINTS; PR00722; CHYMOTRYPSIN.  
PRINTS; PR00261; LDLRECEPTOR.  
SMART; SM00132; LDLa; 3.  
SMART; SM00020; Tryp\_SPC; 1.  
PROSITE; PS01180; CUB; 1.  
PROSITE; PS01209; LDLRA\_1; 1.  
PROSITE; PS00068; LDLRA\_2; 3.  
PROSITE; PS0240; TRYPSIN\_DOM; 1.  
PROSITE; PS00134; TRYPSIN\_HIS; 1.  
PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease; Repeat; Signal-anchor; Transmembrane;  
Glycoprotein; Alternative splicing.  
FT DOMAIN 1 55 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 56 76 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (POTENTIAL).  
FT DOMAIN 77 811 EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 213 336 CUB 1.  
FT DOMAIN 335 452 CUB 2.  
FT DOMAIN 457 489 LDL-RECEPTOR CLASS A 1.  
FT DOMAIN 490 526 LDL-RECEPTOR CLASS A 2.  
FT DOMAIN 530 567 LDL-RECEPTOR CLASS A 3.  
FT DOMAIN 577 811 SERINE PROTEASE.  
FT ACT\_SITE 617 617 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 668 668 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 762 762 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 216 216 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 338 338 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 433 433 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 453 453 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 518 518 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT VARSPLIC 409 461 LGLRILOPVAEIPVAVATAGITINPTSQISLTGPGVRVHY  
FT GLYCNOSDPCEGE -> YHFLSSIMLPFLPPPLSPSTVTP  
FT SLEAQVPLNRGARGARGSGWGCQACP (in isoform 2).  
FT FTID=VSP\_008379.  
FT Missing (in isoform 2).  
FT /FTID=VSP\_008380.  
FT A -> V (IN REF. 4).  
FT 811 AA; 89959 MW; 7E6P193F655DDESD CRC64;  
FT VARSPLIC 462 811  
FT CONFLICT 116 116  
FT SEQUENCE 811 AA; 89959 MW; 7E6P193F655DDESD CRC64;  
Query Match 26.8%; Score 404.5; DB 1; Length 811;  
Best Local Similarity 36.8%; Pred. No. 9.5e-29;  
Matches 99; Conservative 48; Mismatches 89; Indels 33; Gaps 10  
QY 7 ELKFKCGGKTLRRPKFIIGTEFTIENQWFAAIYRRHRGGSVTVVCGSLTSPCWVIS 66  
Db 562 DEHCDGGLQ--GPSRIRVGAVSGEWNQASLQVRGE----HICGALLADRWWIT 614  
QY 67 ATHCFTIDVPKGD-----YIVYLGSRRLNSNTQGMKEVENLILHKDYSDATLAHN 119



615 AAHCF-----QEDSMASVLTWTVFLGKVNQNSRWPGEVFKVSRLLLLHPYHEED--SHDY 667  
120 DIALKIRSKERCAQPSRTIOTICLPMSYNDPFGTSCITG--KENSTDVLYPEQL 177  
668 DVALLOLHPVVR-----SAAVRPCLPARSHFFPGLHCWITGMAUREGGP-----ISNAL 720  
178 KMTVVVQLSHRECQOPHYGSEVTKMLCAADPQKMTDCSQDGGGLVC-SLQGRWTLT 236  
721 QKVDVQLLPQDLCSVRY--QVTPRLMCAVYRKKKACQDGGGLVCALSGRWFLA 778  
237 GIVSGRGCAKDKRGVTVRVSHPLPWR 265  
779 GLVSMGLCGRPNTFVTRITGVISWQ 807  
DB 779 GLVSMGLCGRPNTFVTRITGVISWQ 807  
RESULT 29  
HEPS\_MOUSE  
ID\_HEPS\_MOUSE STANDARD; PRT; 436 AA.  
AC O35453; Q3CW97;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 10-OCT-2003 (Rel. 42, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Serine protease hepsin (EC 3.4.21.-).  
GN HPN.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RC TISSUE=Liver;  
RX MEDLINE=98058912; PubMed=9395459;  
RA Vu T.-X.H., Liu R.W., Haakema C., Tomasek J.J., Howard E.W.;  
RT "Identification and cloning of the membrane-associated serine  
protease, hepsin, from mouse preimplantation embryos.";  
RL J. Biol. Chem. 272:31315-31320(1997).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RX MEDLINE=99339944; PubMed=10411637;  
RA Kawamura S., Kurachi K., Dayashiki Y., Kurachi K.;  
RT "Complete nucleotide sequence, origin of isoform and functional  
characterization of the mouse hepsin gene.";  
RL Eur. J. Biochem. 262:755-764(1999).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC STRAIN=C57BL/6J; TISSUE=Kidney;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Nombaeys P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Yashaw-Boris A., Yoshida K., Haegawa Y., Kawaji H., Kohtsuki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
CC -!- FUNCTION: Plays an essential role in cell growth and maintenance  
of cell morphology.  
CC -!- SUBCELLULAR LOCATION: Type II membrane protein.  
CC -!- ALTERNATIVE PRODUCTS:  
Event-Alternative splicing; Named isoforms=2;  
Name=1; Synonyms=1a;

CC IsoId=O35453-1; Sequence=Displayed;  
CC Note=Minor isoform;  
CC Name=2; Synonyms=2a;  
CC IsoId=O35453-2; Sequence=VSP\_007232;  
CC Note=Major isoform;  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
CC -!- CAUTION: Ref.3 sequence differs from that shown due to  
frameshifts in positions 155, 191 and 233.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce  
or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL; AF030065; AB84221.1; -;  
CC EMBL; AK008694; BAB22289.2; ALT\_FRAME.  
CC HSSP; P00763; IDPO.  
CC MEROPS; S01.224; -.  
CC MGD; MGI:1196620; Hpn.  
CC InterPro; IPR009003; Cys Ser trypsin.  
CC InterPro; IPR001254; Peptidase S1.  
CC InterPro; IPR001314; Peptidase\_S1A.  
CC InterPro; IPR001190; Srrc\_receptor.  
CC Pfam; PF00089; trypsin; 1.  
CC PRINTS; PR00722; CHYMOTRYPSIN.  
CC SMART; SM00202; SR; 1.  
CC SMART; SM00202; TRYPSIN\_SPC; 1.  
CC PROSITE; PS00240; TRYPSIN\_DOM; 1.  
CC PROSITE; PS00334; TRYPSIN\_HIS; 1.  
CC PROSITE; PS00335; TRYPSIN\_SER; 1.  
KW Hydrolyase; Serine protease; Transmembrane; Signal-anchor;  
KW Alternative splicing.  
FT CHAIN 1 181  
FT SERINE PROTEASE HEPsin, NON-CATALYTIC  
FT CHAIN (POTENTIAL).  
FT CHAIN 182 436  
FT SERINE PROTEASE HEPsin, CATALYTIC CHAIN  
FT (POTENTIAL).  
FT DOMAIN 21 36  
FT CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 37 63  
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
FT (POTENTIAL).  
FT DOMAIN 64 436  
FT EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 182 436  
FT SERINE PROTEASE.  
FT ACT\_SITE 222 222  
FT CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 276 276  
FT CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 372 372  
FT CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT DISULFID 172 296  
FT INTERCHAIN (BY SIMILARITY).  
FT DISULFID 207 223  
FT BY SIMILARITY.  
FT DISULFID 341 357  
FT BY SIMILARITY.  
FT DISULFID 368 400  
FT BY SIMILARITY.  
FT CARBOHYD 131 131  
FT N-LINKED (GLCNAC...) (POTENTIAL).  
FT VARSPLIC 25 44  
FT Missing (in isoform 2).  
FT /FTID=VSP\_007232.  
FT L -> F (IN REF. 2 AND 3).  
FT CONFLICT 85 85  
FT L -> Y (IN REF. 3).  
FT CONFLICT 204 204  
FT G -> R (IN REF. 3).  
FT CONFLICT 214 214  
FT NR -> ET (IN REF. 3).  
FT CONFLICT 228 229  
FT P -> L (IN REF. 3).  
FT CONFLICT 264 264  
FT H -> N (IN REF. 3).  
FT CONFLICT 281 281  
SQ SEQUENCE 436 AA; 46787 MW; 4A0993148C620BD0 CRC64;  
Query Match 26.7%; Score 402.5; DB 1; Length 436;  
Best Local Similarity 35.7%; Pred. No. 6.7e-29;  
Matches 101; Conservative 54; Mismatches 85; Indels 43; Gaps 15.  
QY 13 CQKTLRPPKPIIGGEFTIENQPFPAIYRHRGSGVYVCGSLISPCWISATHCFI 72  
DB 172 CQRRKL-PVDRIYVGGQDSLGRWPQVSL--RYDG---THLCGSLSGDWLTAAHCF- 224  
QY 73 DYPKKEDYI-----VYLGRSLNSNTQGENKFEVENLIILH-----KDYSAADTLAHHNDIA 122  
DB 225 --PERNEVLSRWRFVAG--AAVETSPHAVQLGQAVIYHGVLPRDPID--ENSDNIA 278

QY	123	LLKIRSEGRCAQPSRTIQITICLPSMYNDPQFTSCBITGFKENSTDVLYPEQ----	LKN	179
Ddb	279	LVLHSSS-----LPLETYIQPVCLFAAGCALVDGKVCTVTGWGNTQ----FYGQAAMVLOS	330	
QY	180	TWVKLIHSRECCQHYYGVSGEYTTKMLCAAAPQWKTSQCDSDGGPLVC--SLQG--RMTL	235	
Ddb	331	ARVPISNEVCNSPDFYGNQIKPMFCAGYPEGGIDACQSDSGGPFPCDSISGTSRWL	390	
QY	236	TGIVSWGRCALKOKPGVYTRVSHFLPW-----TRSHKYENGIL	274	
Ddb	391	CGIVSWGTCALARKPGVYTKVTDPREMIFRAIKTHS-EASGM	432	
 RESULT 30 KAL_HUMAN STANDARD; PRT; 638 AA.				
ID	KAL_HUMAN	PRT;	638	AA.
AD	P03552;			
DC	23-OCT-1986	(Rel. 02, Created)		
DDT	23-OCT-1986	(Rel. 02, Last sequence update)		
DT	15-MAR-2004	(Rel. 43, Last annotation update)		
DE	Plasma kallikrein precursor (BC 3.4.21.34) (Plasma prekallikrein)			
DE	(kininogenin) (Fletcher factor).			
OS	KLKB1 OR KLK3			
OS	Homo sapiens (Human)			
OC	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=86243359; PubMed=3521732;			
RA	Chung D.W., Fujikawa K., McMullen B.A., Davie E.W.;			
RA	"Human plasma prekallikrein, a zymogen to a serine protease that			
RT	contains four tandem repeats,"			
RT	Biochemistry 25:2410-2417(1986).			
RL	[2]			
RN	SEQUENCE FROM N.A., AND VARIANTS SER-143; GLN-202 AND PRO-208.			
RP	MEDLINE=20487549; PubMed=11031105;			
RA	Yu H., Anderson P.J., Freedman B.I., Rich S.S., Bowden D.W.;			
RA	"Genomic structure of the human plasma prekallikrein gene,			
RT	identification of allelic variants, and analysis in end-stage renal			
RT	disease."			
RL	Genomics 69:225-234 (2000).			
RN	[3]			
RP	SEQUENCE FROM N.A., AND VARIANTS SER-143; THR-178; GLN-202; CYS-269;			
RP	VAL-311; ALA-358; ALA-381; PRO-442 AND GLN-560.			
RA	Rieder M.J., Armet T.Z., Carrington D.P., Ozuna M., Kuldanek S.A.,			
RA	Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;			
RL	Submitted (DEC-2002) to the EMBL/GenBank/DBAJ databases.			
RN	[4]			
RP	PARTIAL SEQUENCE, AND DISULFIDE BONDS.			
RP	MEDLINE=91152016; PubMed=1958666;			
RA	McMullen B.A., Fujikawa K., Davie E.W.;			
RA	"Location of the disulfide bonds in human plasma prekallikrein: the			
RT	presence of four novel apple domains in the amino-terminal portion of			
RT	the molecule."			
RL	Biochemistry 30:2050-2056(1991).			
RN	[5]			
RP	CARBOHYDRATE-LINKAGE SITE ASN-453.			
RX	MEDLINE=22660472; PubMed=14754519;			
RA	Zhang H., Li X.-J., Martin D.B., Aebersold R.;			
RA	"Identification and quantification of N-linked glycoproteins using			
RT	hydrazide chemistry, stable isotopic labeling and mass spectrometry.";			
RT	Nat. Biotechnol. 21:660-666(2003).			
CC	-I- FUNCTION: The enzyme cleaves Lys-Arg and Arg-Ser bonds. It			
CC	activates, in a reciprocal reaction, factor XII after its binding			
CC	to a negatively charged surface. It also releases bradykinin from			
CC	HMW kininogen and may also play a role in the renin-angiotensin			
CC	system by converting prorenin into renin.			
CC	-I- CATALYTIC ACTIVITY: Cleaves selectively Arg- -Xaa and Lys- -Xaa			
CC	bonds, including Lys- -Arg and Arg- -Ser bonds in (human)			
CC	kininogen to release bradykinin.			
CC	-I- SUBUNIT: The zymogen is activated by factor XIa, which cleaves			

DISULFID	137	166
FT		
DISULFID	141	147
FT		
DISULFID	201	284
FT		
DISULFID	227	256
FT		
DISULFID	231	237
FT		
DISULFID	292	375
FT		
DISULFID	318	347
FT		
DISULFID	322	328
FT		
DISULFID	340	345
FT		
DISULFID	383	503
FT		
DISULFID	419	435
FT		
DISULFID	517	584
FT		
DISULFID	548	563
FT		
DISULFID	574	602
FT		
VARIANT	143	143
FT		
VARIANT	178	178
FT		
VARIANT	202	202
FT		
VARIANT	208	208
FT		
VARIANT	269	269
FT		
VARIANT	311	311
FT		
VARIANT	358	358
FT		
VARIANT	381	381
FT		
VARIANT	442	442
FT		
VARIANT	560	560
FT		
SEQUENCE	638 AA; 71369 MW; E62PFC1053838FB4 CRC64;	
SEQ		
Query Match            26.7%; Score 402; DB 1; Length 638;		
Best Local Similarity   33.1%; Pred. No. 1.2e-28;		
Matches   88; Conservative   56; Mismatches   98; Indels   24; Gaps   7;		
QY	17	TLRPFKIKGEFTTIENQPFAIY-----REHRGGSVTVCGGSLISPCWVISATHCF 71
Db	384	TKKSTRIVGTNSWGSEWPQVLSQVKLTAQH-----LCGGSLIQHWLVIAHCF 436
QY	72	IDYPKKEDYIVYLGRSLNNTQGEMKFVENILHKDYSADTLAHNDIALLKIRSKEG 131
Db	437	DGLPLQDVRIYSGILNLSDITKTPSPQIKEIIHQNVKVSEGNH--DIALIKLQAP-- 492
QY	132	RCAQPSRTIQICLPMSYNPQFTSCEIFGF--KENSTDYLYPEOLKMTVVKLISHRE 189
Db	493	--LNTVEFKPICLPFSKGDSTIYTNCWTGWGFSKERGE---IQNILQKNIPLVTNKEE 547
QY	190	CQPHYYGSEVTTKMLCAADPOWKTDSCQDGSGGPLVCSLQGRMILTGVSMGRGCALKD 249
Db	548	CCK-RYQYKIKTQRWCAGYKEGKGDKACKDSGGLPVCKHGMRVLGITSWGECARRE 606
QY	250	KPGVTRVSHFLPWIRSHTEENGLA 275
Db	607	QPGVYTKVAEYMDWILEKTQSSDGKA 632
RESULT 31		
TMSS_MOUSE		
ID	TMSS_MOUSE	STANDARD; PRT; 455 AA.
AC	Q9ER04; Q9ER02; Q9ER03;	
DT	16-OCT-2001 (Rel. 40, Created)	
DT	16-OCT-2001 (Rel. 40, Last sequence update)	
DT	28-FEB-2003 (Rel. 41, Last annotation update)	
DE	Transmembrane protease, serine 5 (EC 3.4.21.-) (Spinesin).	
GN	TMPSR55.	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	

```
FT CARBOHYD 170 170 N-LINKED (GLNAC... ) (POTENTIAL).
FT CARBOHYD 319 319 N-LINKED (GLNAC... ) (POTENTIAL).
FT CARBOHYD 375 375 N-LINKED (GLNAC... ) (POTENTIAL).
FT VARSPLIC 1 144 Missing (in isoform 2).
FT VARSPLIC 1 10 /FtId=VSP 005395.
FT VARSPLIC 1 182 Missing (in isoform 3).
FT VARSPLIC 1 182 /FtId=VSP 005396.
FT VARSPLIC 1 182 Missing (in isoform 1).
FT VARSPLIC 183 192 /FtId=VSP 005397.
FT VARSPLIC 325 325 GGLVEAWKP -> MEAQVGLLMV (in isoform 1).
FT CONFLICT 325 325 /FtId=VSP 005398.
FT CONFLICT 455 AA; 49632 MW; 5CFCG1789C689AA CRC64;
SQ SEQUENCE 26.4k; Score 398.5; DB 1; Length 455;
Best Local Similarity 34.5k; Pred. No. 1.6e-28;
Matches 96; Conservative 42; Mismatches 99; Indels 41; Gaps 8;
Qy 2 PPSPEELK-FQCQKTLRPRFKIIGGEFTIENQFWFAIYRRHGGVYVCGSLIS 60
Db 197 PSGRIVLKCSECCARPLAGR-IVGQAVASGRWQASVMLGSR-----HTCGASVLA 249
Qy 61 PCWVISATHCFIDYPKEDIVVLGRSLNS-----NTQEMKFEVENLIL 106
Db 250 PHWVTAARHCWYGF-----RUSRSSWRVHAGLVSHGAVRQHQTMT---VEKIIP 296
Qy 107 HKDYSADTLAHHNDIALLKIRSKRGCAQPSRTIQTICLPMSYNDPQFGTSCBITGFGXE 166
Db 297 HPLYSAGN-HDYDVALLQLRTP-----INFSDIVDAVCLPAKEQYFPWGSQWVGWHT 350
Qy 167 NSTDYLPEQLKMTVTVKLIHRECOQPHYGVSEVTTMLCAADPQKWTSCQDSDGGPLV 226
Db 351 DPSHTSSDTLQDTMTWELLSTHLCNSSCMYSGALTHRLMCAGYLDGRADACQDSDGGPLV 410
Qy 227 CSLQGRMTLGIWVGWGGCALKDKPGVYTRVSHFLPMI 264
Db 411 CFSGDTHLVGVVSWGSGCAEPNRPVGYAVAFELDWI 448
RESULT 32
HEPS HUMAN
ID HEPHS HUMAN STANDARD; PRT; 417 AA.
AC P05981;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serine protease hepsin (EC 3.4.21.-) (Transmembrane protease, serine
DE 1).
GN HPN OR TMFRSS1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=88209431; PubMed=2835076;
RA Levitus S.P., Loeb K.R., Hagen F.S., Kurachi K., Davie E.W.;
RT "A novel trypsin-like serine protease (hepsin) with a putative
RT transmembrane domain expressed by human liver and hepatoma cells.";
RL Biochemistry 27:1067-1074 (1988).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas, and Spleen;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins L.F., Jordan B., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Tothiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
```

```
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fabey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grawood J., Schmutz J., Myers R.M.,
RA Butterfield J.S.N., Krzywinski W.I., Skalska U., Smalhus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=91358502; PubMed=1885621;
RA Tsuji A., Torres-Rosado A., Aral T., le Beau M.M., Lemons R.S.,
RA Chou S.H., Kurachi K.;
RT "Hepsin, a cell membrane-associated protease. Characterization,
RT tissue distribution, and gene localization.";
RL J. Biol. Chem. 266:16948-16953 (1991).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=93348237; PubMed=8346233;
RA Torres-Rosado A., O'Shea K.S., Tsuji A., Chou S.H., Kurachi K.;
RT "Hepsin, a putative cell-surface serine protease, is required for
RT mammalian cell growth.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:7181-7187 (1993).
CC -!- FUNCTION: Plays an essential role in cell growth and maintenance
CC of cell morphology.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC -!- TISSUE SPECIFICITY: Present in most tissues, with the highest
CC level in liver.
CC -!- SIMILARITY: Belongs to peptidase family S1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboratio
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on it
CC use by non-profit institutions as long as its content is in no wa
CC modified and this statement is not removed. Usage by and for commercia
CC entities requires a license agreement (see http://www.isb-sib.ch/announce
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M18930; AAA36013.1; -
CC EMBL; X07732; CAA30558.1; -
CC EMBL; X07002; CAA30058.1; -
CC EMBL; BC025716; AAB25716.1; -
CC F1R; S00845; S00845.
CC HSP; P00763; LDPO.
CC MEROPS; S01.224; -.
CC Genew; HGNC:5155; HPN.
CC MIM; 142440; -.
CC GO; GO:0005887; C: integral to plasma membrane; TAS.
CC GO; GO:0008236; F: serine-type peptidase activity; TAS.
CC GO; GO:0008151; P: cell growth and/or maintenance; TAS.
CC InterPro; IPR009003; Cys Ser trypsin.
CC InterPro; IPR001254; Peptidase S1.
CC InterPro; IPR001314; Peptidase_S1A.
CC Pfam; PF00089; trypsin; 1
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC SMART; SM00020; Tryp_Spc; 1.
CC PROSITE; PS0240; TRYPsin DOM; 1.
CC PROSITE; PS00134; TRYPsin HIS; 1.
CC PROSITE; PS00135; TRYPsin SER; 1.
CC Hydrolase; Serine protease; Transmembrane; Signal-anchor.
CX CHAIN 1 162 SERINE PROTEASE HEPsin, NON-CATALYTIC
FT CHAIN 163 417 SERINE PROTEASE HEPsin, CATALYTIC CHAIN
(POTENTIAL).
FT DOMAIN 1 17 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 18 44 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
FT DOMAIN 45 417 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 163 417 SERINE PROTEASE.
```

HSP;	P00763; IDPO.
MEROPS;	S01.151; ..
InterPro;	IPR009003; Cys_Ser trypsin.
InterPro;	IPR001254; Peptidase S1.
InterPro;	IPR001314; Peptidase_S1A.
Pfam;	PF00089; trypsin; 1.
PRINTS;	PR00722; CHYMOTRYPSIN.
SMART;	SM00020; Tryp_Spc; 1.
PROSITE;	PS00240; TRYPSIN DOM; 1.
PROSITE;	PS00134; TRYPSIN_HIS; 1.
PROSITE;	PS00135; TRYPSIN_SER; 1.
Hydrolase;	Serine protease; Digestion; Zymogen;
Cation-binding;	Signal; Multigene family.
SIGNAL	1 16 BY SIMILARITY.
PROPEP	17 25 ACTIVATION PEPTIDE (BY SIMILARITY).
CHAIN	26 248 TRYP SIN II-P25.
ACT SITE	65 65 CHARGE RELAY SYSTEM (BY SIMILARITY).
METAL	77 77 CALCIUM (BY SIMILARITY).
METAL	79 79 CALCIUM (VIA CARBONYL OXYGEN)
	(BY SIMILARITY).
METAL	82 82 CALCIUM (VIA CARBONYL OXYGEN)
	(BY SIMILARITY).
METAL	87 87 CALCIUM (BY SIMILARITY).
ACT SITE	109 109 CHARGE RELAY SYSTEM (BY SIMILARITY).
ACT SITE	202 202 CHARGE RELAY SYSTEM (BY SIMILARITY).
DLSULFD	32 162 BY SIMILARITY.
DLSULFD	50 66 BY SIMILARITY.
DLSULFD	134 235 BY SIMILARITY.
DLSULFD	141 208 BY SIMILARITY.
DLSULFD	173 187 BY SIMILARITY.
DLSULFD	198 222 BY SIMILARITY.
SITE	196 196 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
SEQUENCE	248 AA; 25622 MW; ESE16B07622B2858E CRC64;

[illegible]

RP SEQUENCE FROM N.A.  
RC TISSUE-Pancreas;  
RA Okajima T., Maniwa M., Nagao S., Fujikawa H., Goto S.;  
RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RX SEQUENCE OF 15-243, AND DISULFIDE BONDS.  
RY MEDLINE=67168848; PubMed=5967094;  
RA Mikes O., Holeysovsky V., Tomasek V., Sorm F.;  
RT "Covalent structure of bovine trypsinogen. The position of the  
RL remaining amides";  
RN Biochem. Biophys. Res. Commun. 24:346-352(1966).  
RN [3]  
RP REVISIONS.  
RX MEDLINE=72035053; PubMed=4399051;  
RA Hartley B.S.;  
RT "Homologies in serine proteinases";  
RN Phillos. Trans. R. Soc. Lond., B, Biol. Sci. 257:77-87(1970).  
RN [4]  
RP REVISIONS.  
RX MEDLINE=75146445; PubMed=1092332;  
RA Titani K., Ericsson L.H., Neurath H., Walsh K.A.;  
RT "Amino acid sequence of dogfish trypsin";  
RN Biochemistry 14:1358-1366(1975).  
RN [5]  
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF CALCIUM-BINDING SITE.  
RX MEDLINE=76072097; PubMed=512;  
RA Bode W., Schwager P.;  
RT "The refined crystal structure of bovine beta-trypsin at 1.8-A  
RL resolution. II. Crystallographic refinement, calcium binding site,  
RN benzamide binding site and active site at pH 7.0.";  
J. Mol. Biol. 98:693-717(1975).  
RN [6]  
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
RX MEDLINE=77112431; PubMed=556951;  
RA Kossiakoff A.A., Chambers J.L., Kay L.M., Stroud R.M.;  
RT "Structure of bovine trypsinogen at 1.9-A resolution.";  
RN Biochemistry 16:654-664(1977).  
RN [7]  
RP DISULFIDE BONDS.  
RX MEDLINE=66079271; PubMed=5892911;  
RA Kauffman D.L.;  
RT "The disulphide bridges of trypsin.";  
RN J. Mol. Biol. 12:929-932(1965).  
RN [8]  
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.  
CC -1- COFACTOR: Binds 1 calcium ion per subunit.  
CC -1- SUBCELLULAR LOCATION: Extracellular.  
CC -1- TISSUE SPECIFICITY: Synthesized in the acinar cells of the  
RN pancreas.  
CC -1- PTM: AUTOCATALYTIC CLEAVAGE AFTER LYS-20 LEADS TO BETA-TRYPSIN BY  
RN RELEASING A TERMINAL HEXAPEPTIDE. SUBSEQUENT CLEAVAGE AFTER  
CC LYS-145 LEADS TO ALPHA-TRYPSIN. FURTHER CLEAVAGE AFTER LYS-190  
CC YIELDS PSEUDOTRYPSIN. A CLEAVAGE MAY ALSO OCCUR AFTER ARG-119.  
CC -1- SIMILARITY: Belongs to peptidase family S1.  
CC -1- DATABASE: NAME=worthington enzyme manual.  
RN WWW="http://www.worthington-biochem.com/TRY/".  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: D38507; BAA07516.1; -;  
DR PDB: 1A07; 25-FEB-98.  
DR PDB: 1AUJ; 14-OCT-98.  
DR PDB: 1A28; 13-JAN-99.  
DR PDB: 1BJU; 13-JAN-99.  
DR PDB: 1BJV; 13-JAN-99.  
DR PDB: 1BTJ; 29-JAN-96.  
DR PDB: 1BTW; 15-OCT-95.  
DR PDB: 1BTX; 15-OCT-95.  
DR PDB: 1BTY; 15-OCT-95.  
DR PDB: 1BZ; 15-OCT-95.  
DR PDB: 1CIN; 30-JAN-02.  
DR PDB: 1CJO; 26-SEP-01.  
DR PDB: 1CJP; 26-SEP-01.  
DR PDB: 1CIQ; 26-SEP-01.  
DR PDB: 1CIR; 26-SEP-01.  
DR PDB: 1CIS; 26-JUL-00.  
DR PDB: 1CIT; 26-SEP-01.  
DR PDB: 1CJD; 26-SEP-01.  
DR PDB: 1CJE; 26-SEP-01.  
DR PDB: 1CJF; 26-SEP-01.  
DR PDB: 1CJG; 26-SEP-01.  
DR PDB: 1CJH; 26-SEP-01.  
DR PDB: 1CJI; 26-SEP-01.  
DR PDB: 1CJL; 26-SEP-01.  
DR PDB: 1CJM; 26-SEP-01.  
DR PDB: 1CJN; 26-SEP-01.  
DR PDB: 1CJO; 26-SEP-01.  
DR PDB: 1CJP; 26-SEP-01.  
DR PDB: 1CJQ; 26-SEP-01.  
DR PDB: 1CJR; 26-SEP-01.  
DR PDB: 1CJS; 26-SEP-01.  
DR PDB: 1CJT; 26-SEP-01.  
DR PDB: 1CJU; 26-SEP-01.  
DR PDB: 1CJV; 26-SEP-01.  
DR PDB: 1CJW; 26-SEP-01.  
DR PDB: 1CJX; 26-SEP-01.  
DR PDB: 1CJY; 26-SEP-01.  
DR PDB: 1CJZ; 26-SEP-01.  
DR PDB: 1CJA; 26-SEP-01.  
DR PDB: 1CJB; 26-SEP-01.  
DR PDB: 1CJC; 26-SEP-01.  
DR PDB: 1CJD; 26-SEP-01.  
DR PDB: 1CJE; 26-SEP-01.  
DR PDB: 1CJF; 26-SEP-01.  
DR PDB: 1CJG; 26-SEP-01.  
DR PDB: 1CJH; 26-SEP-01.  
DR PDB: 1CJI; 26-SEP-01.  
DR PDB: 1CJL; 26-SEP-01.  
DR PDB: 1CJM; 26-SEP-01.  
DR PDB: 1CJN; 26-SEP-01.  
DR PDB: 1CJO; 26-SEP-01.  
DR PDB: 1CJP; 26-SEP-01.  
DR PDB: 1CJQ; 26-SEP-01.  
DR PDB: 1CJR; 26-SEP-01.  
DR PDB: 1CJS; 26-SEP-01.  
DR PDB: 1CJT; 26-SEP-01.  
DR PDB: 1CJU; 26-SEP-01.  
DR PDB: 1CJV; 26-SEP-01.  
DR PDB: 1CJW; 26-SEP-01.  
DR PDB: 1CJX; 26-SEP-01.  
DR PDB: 1CJY; 26-SEP-01.  
DR PDB: 1CJZ; 26-SEP-01.  
DR PDB: 1CJA; 26-SEP-01.  
DR PDB: 1CJB; 26-SEP-01.  
DR PDB: 1CJC; 26-SEP-01.  
DR PDB: 1CJD; 26-SEP-01.  
DR PDB: 1CJE; 26-SEP-01.  
DR PDB: 1CJF; 26-SEP-01.  
DR PDB: 1CJG; 26-SEP-01.  
DR PDB: 1CJH; 26-SEP-01.  
DR PDB: 1CJI; 26-SEP-01.  
DR PDB: 1CJL; 26-SEP-01.  
DR PDB: 1CJM; 26-SEP-01.  
DR PDB: 1CJN; 26-SEP-01.  
DR PDB: 1CJO; 26-SEP-01.  
DR PDB: 1CJP; 26-SEP-01.  
DR PDB: 1CJQ; 26-SEP-01.  
DR PDB: 1CJR; 26-SEP-01.  
DR PDB: 1CJS; 26-SEP-01.  
DR PDB: 1CJT; 26-SEP-01.  
DR PDB: 1CJU; 26-SEP-01.  
DR PDB: 1CJV; 26-SEP-01.  
DR PDB: 1CJW; 26-SEP-01.  
DR PDB: 1CJX; 26-SEP-01.  
DR PDB: 1CJY; 26-SEP-01.  
DR PDB: 1CJZ; 26-SEP-01.  
DR PDB: 1CJA; 26-SEP-01.  
DR PDB: 1CJB; 26-SEP-01.  
DR PDB: 1CJC; 26-SEP-01.  
DR PDB: 1CJD; 26-SEP-01.  
DR PDB: 1CJE; 26-SEP-01.  
DR PDB: 1CJF; 26-SEP-01.  
DR PDB: 1CJG; 26-SEP-01.  
DR PDB: 1CJH; 26-SEP-01.  
DR PDB: 1CJI; 26-SEP-01.  
DR PDB: 1CJL; 26-SEP-01.  
DR PDB: 1CJM; 26-SEP-01.  
DR PDB: 1CJN; 26-SEP-01.  
DR PDB: 1CJO; 26-SEP-01.  
DR PDB: 1CJP; 26-SEP-01.  
DR PDB: 1CJQ; 26-SEP-01.  
DR PDB: 1CJR; 26-SEP-01.  
DR PDB: 1CJS; 26-SEP-01.  
DR PDB: 1CJT; 26-SEP-01.  
DR PDB: 1CJU; 26-SEP-01.  
DR PDB: 1CJV; 26-SEP-01.  
DR PDB: 1CJW; 26-SEP-01.  
DR PDB: 1CJX; 26-SEP-01.  
DR PDB: 1CJY; 26-SEP-01.  
DR PDB: 1CJZ; 26-SEP-01.  
DR PDB: 1CJA; 26-SEP-01.  
DR PDB: 1CJB; 26-SEP-01.  
DR PDB: 1CJC; 26-SEP-01.  
DR PDB: 1CJD; 26-SEP-01.  
DR PDB: 1CJE; 26-SEP-01.  
DR PDB: 1CJF; 26-SEP-01.  
DR PDB: 1CJG; 26-SEP-01.  
DR PDB: 1CJH; 26-SEP-01.  
DR PDB: 1CJI; 26-SEP-01.  
DR PDB: 1CJL; 26-SEP-01.  
DR PDB: 1CJM; 26-SEP-01.  
DR PDB: 1CJN; 26-SEP-01.  
DR PDB: 1CJO; 26-SEP-01.  
DR PDB: 1CJP; 26-SEP-01.  
DR PDB: 1CJQ; 26-SEP-01.  
DR PDB: 1CJR; 26-SEP-01.  
DR PDB: 1CJS; 26-SEP-01.  
DR PDB: 1CJT; 26-SEP-01.  
DR PDB: 1CJU; 26-SEP-01.  
DR PDB: 1CJV; 26-SEP-01.  
DR PDB: 1CJW; 26-SEP-01.  
DR PDB: 1CJX; 26-SEP-01.  
DR PDB: 1CJY; 26-SEP-01.  
DR PDB: 1CJZ; 26-SEP-01.  
DR PDB: 1CJA; 26-SEP-01.  
DR PDB: 1CJB; 26-SEP-01.  
DR PDB: 1CJC; 26-SEP-01.  
DR PDB: 1CJD; 26-SEP-01.  
DR PDB: 1CJE; 26-SEP-01.  
DR PDB: 1CJF; 26-SEP-01.  
DR PDB: 1CJG; 26-SEP-01.  
DR PDB: 1CJH; 26-SEP-01.  
DR PDB: 1CJI; 26-SEP-01.  
DR PDB: 1CJL; 26-SEP-01.  
DR PDB: 1CJM; 26-SEP-01.  
DR PDB: 1CJN; 26-SEP-01.  
DR PDB: 1CJO; 26-SEP-01.  
DR PDB: 1CJP; 26-SEP-01.  
DR PDB: 1CJQ; 26-SEP-01.  
DR PDB: 1CJR; 26-SEP-01.  
DR PDB: 1CJS; 26-SEP-01.  
DR PDB: 1CJT; 26-SEP-01.  
DR PDB: 1CJU; 26-SEP-01.  
DR PDB: 1CJV; 26-SEP-01.  
DR PDB: 1CJW; 26-SEP-01.  
DR PDB: 1CJX; 26-SEP-01.  
DR PDB: 1CJY; 26-SEP-01.  
DR PDB: 1CJZ; 26-SEP-01.  
DR PDB: 1CJA; 26-SEP-01.  
DR PDB: 1CJB; 26-SEP-01.  
DR PDB: 1CJC; 26-SEP-01.  
DR PDB: 1CJD; 26-SEP-01.  
DR PDB: 1CJE; 26-SEP-01.  
DR PDB: 1CJF; 26-SEP-01.  
DR PDB: 1CJG; 26-SEP-01.  
DR PDB: 1CJH; 26-SEP-01.  
DR PDB: 1CJI; 26-SEP-01.  
DR PDB: 1CJL; 26-SEP-01.  
DR PDB: 1CJM; 26-SEP-01.  
DR PDB: 1CJN; 26-SEP-01.  
DR PDB: 1CJO; 26-SEP-01.  
DR PDB: 1CJP; 26-SEP-01.  
DR PDB: 1CJQ; 26-SEP-01.  
DR PDB: 1CJR; 26-SEP-01.  
DR PDB: 1CJS; 26-SEP-01.  
DR PDB: 1CJT; 26-SEP-01.  
DR PDB: 1CJU; 26-SEP-01.  
DR PDB: 1CJV; 26-SEP-01.  
DR PDB: 1CJW; 26-SEP-01.  
DR PDB: 1CJX; 26-SEP-01.  
DR PDB: 1CJY; 26-SEP-01.  
DR PDB: 1CJZ; 26-SEP-01.  
DR PDB: 1CJA; 26-SEP-01.  
DR PDB: 1CJB; 26-SEP-01.  
DR PDB: 1CJC; 26-SEP-01.  
DR PDB: 1CJD; 26-SEP-01.  
DR PDB: 1CJE; 26-SEP-01.  
DR PDB: 1CJF; 26-SEP-01.  
DR PDB: 1CJG; 26-SEP-01.  
DR PDB: 1CJH; 26-SEP-01.  
DR PDB: 1CJI; 26-SEP-01.  
DR PDB: 1CJL; 26-SEP-01.  
DR PDB: 1CJM; 26-SEP-01.  
DR PDB: 1CJN; 26-SEP-01.  
DR PDB: 1CJO; 26-SEP-01.  
DR PDB: 1CJP; 26-SEP-01.  
DR PDB: 1CJQ; 26-SEP-01.  
DR PDB: 1CJR; 26-SEP-01.  
DR PDB: 1CJS; 26-SEP-01.  
DR PDB: 1CJT; 26-SEP-01.  
DR PDB: 1CJU; 26-SEP-01.  
DR PDB: 1CJV; 26-SEP-01.  
DR PDB: 1CJW; 26-SEP-01.  
DR PDB: 1CJX; 26-SEP-01.  
DR PDB: 1CJY; 26-SEP-01.  
DR PDB: 1CJZ; 26-SEP-01.  
DR PDB: 1CJA; 26-SEP-01.  
DR PDB: 1CJB; 26-SEP-01.  
DR PDB: 1CJC; 26-SEP-01.  
DR PDB: 1CJD; 26-SEP-01.  
DR PDB: 1CJE; 26-SEP-01.  
DR PDB: 1CJF; 26-SEP-01.  
DR PDB: 1CJG; 26-SEP-01.  
DR PDB: 1CJH; 26-SEP-01.  
DR PDB: 1CJI; 26-SEP-01.  
DR PDB: 1CJL; 26-SEP-01.  
DR PDB: 1CJM; 26-SEP-01.  
DR PDB: 1CJN; 26-SEP-01.  
DR PDB: 1CJO; 26-SEP-01.  
DR PDB: 1CJP; 26-SEP-01.  
DR PDB: 1CJQ; 26-SEP-01.  
DR PDB: 1CJR; 26-SEP-01.  
DR PDB: 1CJS; 26-SEP-01.  
DR PDB: 1CJT; 26-SEP-01.  
DR PDB: 1CJU; 26-SEP-01.  
DR PDB: 1CJV; 26-SEP-01.  
DR PDB: 1CJW; 26-SEP-01.  
DR PDB: 1CJX; 26-SEP-01.  
DR PDB: 1CJY; 26-SEP-01.  
DR PDB: 1CJZ; 26-SEP-01.  
DR PDB: 1CJA; 26-SEP-01.  
DR PDB: 1CJB; 26-SEP-01.  
DR PDB: 1CJC; 26-SEP-01.  
DR PDB: 1CJD; 26-SEP-01.  
DR PDB: 1CJE; 26-SEP-01.  
DR PDB: 1CJF; 26-SEP-01.  
DR PDB: 1CJG; 26-SEP-01.  
DR PDB: 1CJH; 26-SEP-01.  
DR PDB: 1CJI; 26-SEP-01.  
DR PDB: 1CJL; 26-SEP-01.  
DR PDB: 1CJM; 26-SEP-01.  
DR PDB: 1CJN; 26-SEP-01.  
DR PDB: 1CJO; 26-SEP-01.  
DR PDB: 1CJP; 26-SEP-01.  
DR PDB: 1CJQ; 26-SEP-01.  
DR PDB: 1CJR; 26-SEP-01.  
DR PDB: 1CJS; 26-SEP-01.  
DR PDB: 1CJT; 26-SEP-01.  
DR PDB: 1CJU; 26-SEP-01.  
DR PDB: 1CJV; 26-SEP-01.  
DR PDB: 1CJW; 26-SEP-01.  
DR PDB: 1CJX; 26-SEP-01.  
DR PDB: 1CJY; 26-SEP-01.  
DR PDB: 1CJZ; 26-SEP-01.  
DR PDB: 1CJA; 26-SEP-01.  
DR PDB: 1CJB; 26-SEP-01.  
DR PDB: 1CJC; 26-SEP-01.  
DR PDB: 1CJD; 26-SEP-01.  
DR PDB: 1CJE; 26-SEP-01.  
DR PDB: 1CJF; 26-SEP-01.  
DR PDB: 1CJG; 26-SEP-01.  
DR PDB: 1CJH; 26-SEP-01.  
DR PDB: 1CJI; 26-SEP-01.  
DR PDB: 1CJL; 26-SEP-01.  
DR PDB: 1CJM; 26-SEP-01.  
DR PDB: 1CJN; 26-SEP-01.  
DR PDB: 1CJO; 26-SEP-01.  
DR PDB: 1CJP; 26-SEP-01.  
DR PDB: 1CJQ; 26-SEP-01.  
DR PDB: 1CJR; 26-SEP-01.  
DR PDB: 1CJS; 26-SEP-01.  
DR PDB: 1CJT; 26-SEP-01.  
DR PDB: 1CJU; 26-SEP-01.  
DR PDB: 1CJV; 26-SEP-01.  
DR PDB: 1CJW; 26-SEP-01.  
DR PDB: 1CJX; 26-SEP-01.  
DR PDB: 1CJY; 26-SEP-01.  
DR PDB: 1CJZ; 26-SEP-01.  
DR PDB: 1CJA; 26-SEP-01.  
DR PDB: 1CJB; 26-SEP-01.  
DR PDB: 1CJC; 26-SEP-01.  
DR PDB: 1CJD; 26-SEP-01.  
DR PDB: 1CJE; 26-SEP-01.  
DR PDB: 1CJF; 26-SEP-01.  
DR PDB: 1CJG; 26-SEP-01.  
DR PDB: 1CJH; 26-SEP-01.  
DR PDB: 1CJI; 26-SEP-01.  
DR PDB: 1CJL; 26-SEP-01.  
DR PDB: 1CJM; 26-SEP-01.  
DR PDB: 1CJN; 26-SEP-01.  
DR PDB: 1CJO; 26-SEP-01.  
DR PDB: 1CJP; 26-SEP-01.  
DR PDB: 1CJQ; 26-SEP-01.  
DR PDB: 1CJR; 26-SEP-01.  
DR PDB: 1CJS; 26-SEP-01.  
DR PDB: 1CJT; 26-SEP-01.  
DR PDB: 1CJU; 26-SEP-01.  
DR PDB: 1CJV; 26-SEP-01.  
DR PDB: 1CJW; 26-SEP-01.  
DR PDB: 1CJX; 26-SEP-01.  
DR PDB: 1CJY; 26-SEP-01.  
DR PDB: 1CJZ; 26-SEP-01.  
DR PDB: 1CJA; 26-SEP-01.  
DR PDB: 1CJB; 26-SEP-01.  
DR PDB: 1CJC; 26-SEP-01.  
DR PDB: 1CJD; 26-SEP-01.  
DR PDB: 1CJE; 26-SEP-01.  
DR PDB: 1CJF; 26-SEP-01.  
DR PDB: 1CJG; 26-SEP-01.  
DR PDB: 1CJH; 26-SEP-01.  
DR PDB: 1CJI; 26-SEP-01.  
DR PDB: 1CJL; 26-SEP-01.  
DR PDB: 1CJM; 26-SEP-01.  
DR PDB: 1CJN; 26-SEP-01.  
DR PDB: 1CJO; 26-SEP-01.  
DR PDB: 1CJP; 26-SEP-01.  
DR PDB: 1CJQ; 26-SEP-01.  
DR PDB: 1CJR; 26-SEP-01.  
DR PDB: 1CJS; 26-SEP-01.  
DR PDB: 1CJT; 26-SEP-01.  
DR PDB: 1CJU; 26-SEP-01.  
DR PDB: 1CJV; 26-SEP-01.  
DR PDB: 1CJW; 26-SEP-01.  
DR PDB: 1CJX; 26-SEP-01.  
DR PDB: 1CJY; 26-SEP-01.  
DR PDB: 1CJZ; 26-SEP-01.  
DR PDB: 1CJA; 26-SEP-01.  
DR PDB: 1CJB; 26-SEP-01.  
DR PDB: 1CJC; 26-SEP-01.  
DR PDB: 1CJD; 26-SEP-01.  
DR PDB: 1CJE; 26-SEP-01.  
DR PDB: 1CJF; 26-SEP-01.  
DR PDB: 1CJG; 26-SEP-01.  
DR PDB: 1CJH; 26-SEP-01.  
DR PDB: 1CJI; 26-SEP-01.  
DR PDB: 1CJL; 26-SEP-01.  
DR PDB: 1CJM; 26-SEP-01.  
DR PDB: 1CJN; 26-SEP-01.  
DR PDB: 1CJO; 26-SEP-01.  
DR PDB: 1CJP; 26-SEP-01.  
DR PDB: 1CJQ; 26-SEP-01.  
DR PDB: 1CJR; 26-SEP-01.  
DR PDB: 1CJS; 26-SEP-01.  
DR PDB: 1CJT; 26-SEP-01.  
DR PDB: 1CJU; 26-SEP-01.  
DR PDB: 1CJV; 26-SEP-01.  
DR PDB: 1CJW; 26-SEP-01.  
DR PDB: 1CJX; 26-SEP-01.  
DR PDB: 1CJY; 26-SEP-01.  
DR PDB: 1CJZ; 26-SEP-01.  
DR PDB: 1CJA; 26-SEP-01.  
DR PDB: 1CJB; 26-SEP-01.  
DR PDB: 1CJC; 26-SEP-01.  
DR PDB: 1CJD; 26-SEP-01.  
DR PDB: 1CJE; 26-SEP-01.  
DR PDB: 1CJF; 26-SEP-01.  
DR PDB: 1CJG; 26-SEP-01.  
DR PDB: 1CJH; 26-SEP-01.  
DR PDB: 1CJI; 26-SEP-01.  
DR PDB: 1CJL; 26-SEP-01.  
DR PDB: 1CJM; 26-SEP-01.  
DR PDB: 1CJN; 26-SEP-01.  
DR PDB: 1CJO; 26-SEP-01.  
DR PDB: 1CJP; 26-SEP-01.  
DR PDB: 1CJQ; 26-SEP-01.  
DR PDB: 1CJR; 26-SEP-01.  
DR PDB: 1CJS; 26-SEP-01.  
DR PDB: 1CJT; 26-SEP-01.  
DR PDB: 1CJU; 26-SEP-01.  
DR PDB: 1CJV; 26-SEP-01.  
DR PDB: 1CJW; 26-SEP-01.  
DR PDB: 1CJX; 26-SEP-01.  
DR PDB: 1CJY; 26-SEP-01.  
DR PDB: 1CJZ; 26-SEP-01.  
DR PDB: 1CJA; 26-SEP-01.  
DR PDB: 1CJB; 26-SEP-01.  
DR PDB: 1CJC; 26-SEP-01.  
DR PDB: 1CJD; 26-SEP-01.  
DR PDB: 1CJE; 26-SEP-01.  
DR PDB: 1CJF; 26-SEP-01.  
DR PDB: 1CJG; 26-SEP-01.  
DR PDB: 1CJH; 26-SEP-01.  
DR PDB: 1CJI; 26-SEP-01.  
DR PDB: 1CJL; 26-SEP-01.  
DR PDB: 1CJM; 26-SEP-01.  
DR PDB: 1CJN; 26-SEP-01.  
DR PDB: 1CJO; 26-SEP-01.  
DR PDB: 1CJP; 26-SEP-01.  
DR PDB: 1CJQ; 26-SEP-01.  
DR PDB: 1CJR; 26-SEP-01.  
DR PDB: 1CJS; 26-SEP-01.  
DR PDB: 1CJT; 26-SEP-01.  
DR PDB: 1CJU; 26-SEP-01.  
DR PDB: 1CJV; 26-SEP-01.  
DR PDB: 1CJW; 26-SEP-01.  
DR PDB: 1CJX; 26-SEP-01.  
DR PDB: 1CJY; 26-SEP-01.  
DR PDB: 1CJZ; 26-SEP-01.  
DR PDB: 1CJA; 26-SEP-01.  
DR PDB: 1CJB; 26-SEP-01.  
DR PDB: 1CJC; 26-SEP-01.  
DR PDB: 1CJD; 26-SEP-01.  
DR PDB: 1CJE; 26-SEP-01.  
DR PDB: 1CJF; 26-SEP-01.  
DR PDB: 1CJG; 26-SEP-01.  
DR PDB: 1CJH; 26-SEP-01.  
DR PDB: 1CJI; 26-SEP-01.  
DR PDB: 1CJL; 26-SEP-01.  
DR PDB: 1CJM; 26-SEP-01.  
DR PDB: 1CJN; 26-SEP-01.  
DR PDB: 1CJO; 26-SEP-01.  
DR PDB: 1CJP; 26-SEP-01.  
DR PDB: 1CJQ; 26-SEP-01.  
DR PDB: 1CJR; 26-SEP-01.  
DR PDB: 1CJS; 26-SEP-01.  
DR PDB: 1CJT; 26-SEP-01.  
DR PDB: 1CJU; 26-SEP-01.  
DR PDB: 1CJV; 26-SEP-01.  
DR PDB: 1CJW; 26-SEP-01.  
DR PDB: 1CJX; 26-SEP-01.  
DR PDB: 1CJY; 26-SEP-01.  
DR PDB: 1CJZ; 26-SEP-01.  
DR PDB: 1CJA; 26-SEP-01.  
DR PDB: 1CJB; 26-SEP-01.  
DR PDB: 1CJC; 26-SEP-01.  
DR PDB: 1CJD; 26-SEP-01.  
DR PDB: 1CJE; 26-SEP-01.  
DR PDB: 1CJF; 26-SEP-01.  
DR PDB: 1CJG; 26-SEP-01.  
DR PDB: 1CJH; 26-SEP-01.  
DR PDB: 1CJI; 26-SEP-01.  
DR PDB: 1CJL; 26-SEP-01.  
DR PDB: 1CJM; 26-SEP-01.  
DR PDB: 1CJN; 26-SEP-01.  
DR PDB: 1CJO; 26-SEP-01.  
DR PDB: 1CJP; 26-SEP-01.  
DR PDB: 1CJQ; 26-SEP-01.  
DR PDB: 1CJR; 26-SEP-01.  
DR PDB: 1CJS; 26-SEP-01.  
DR PDB: 1CJT; 26-SEP-01.  
DR PDB: 1CJU; 26-SEP-01.  
DR PDB: 1CJV; 26-SEP-01.  
DR PDB: 1CJW; 26-SEP-01.  
DR PDB: 1CJX; 26-SEP-01.  
DR PDB: 1CJY; 26-SEP-01.  
DR PDB: 1CJZ; 26-SEP-01.  
DR PDB: 1CJA; 26-SEP-01.  
DR PDB: 1CJB; 26-SEP-01.  
DR PDB: 1CJC; 26-SEP-01.  
DR PDB: 1CJD; 26-SEP-01.  
DR PDB: 1CJE; 26-SEP-01.  
DR PDB: 1CJF; 26-SEP-01.  
DR PDB: 1CJG; 26-SEP-01.  
DR PDB: 1CJH; 26-SEP-01.  
DR PDB: 1CJI; 26-SEP-01.  
DR PDB: 1CJL; 26-SEP-01.  
DR PDB: 1CJM; 26-SEP-01.  
DR PDB: 1CJN; 26-SEP-01.  
DR PDB: 1CJO; 26-SEP-01.  
DR PDB: 1CJP; 26-SEP-01.  
DR PDB: 1CJQ; 26-SEP-01.  
DR PDB: 1CJR; 26-SEP-01.  
DR PDB: 1CJS; 26-SEP-01.  
DR PDB: 1CJT; 26-SEP-01.  
DR PDB: 1CJU; 26-SEP-01.  
DR PDB: 1CJV; 26-SEP-01.  
DR PDB: 1CJW; 26-SEP-01.  
DR PDB: 1CJX; 26-SEP-01.  
DR PDB: 1CJY; 26-SEP-01.  
DR PDB: 1CJZ; 26-SEP-01.  
DR PDB: 1CJA; 26-SEP-01.  
DR PDB: 1CJB; 26-SEP-01.  
DR PDB: 1CJC; 26-SEP-01.  
DR PDB: 1CJD; 26-SEP-01.  
DR PDB: 1CJE; 26-SEP-01.  
DR PDB: 1CJF; 26-SEP-01.  
DR PDB: 1CJG; 26-SEP-01.  
DR PDB: 1CJH; 26-SEP-01.  
DR PDB: 1CJI; 26-SEP-01.  
DR PDB: 1CJL; 26-SEP-01.  
DR PDB: 1CJM; 26-SEP-01.  
DR PDB: 1CJN; 26-SEP-01.  
DR PDB: 1CJO; 26-SEP-01.  
DR PDB: 1CJP; 26-SEP-01.  
DR PDB: 1CJQ; 26-SEP-01.  
DR PDB: 1CJR; 26-SEP-01.  
DR PDB: 1CJS; 26-SEP-01.  
DR PDB: 1CJT; 26-SEP-01.  
DR PDB: 1CJU; 26-SEP-01.  
DR PDB: 1CJV; 26-SEP-01.  
DR PDB: 1CJW; 26-SEP-01.  
DR PDB: 1CJX; 26-SEP-01.  
DR PDB: 1CJY; 26-SEP-01.  
DR PDB: 1CJZ; 26-SEP-01.  
DR PDB: 1CJA; 26-SEP-01.  
DR PDB: 1CJB; 26-SEP-01.  
DR PDB: 1CJC; 26-SEP-01.  
DR PDB: 1CJD; 26-SEP-01.  
DR PDB: 1CJE; 26-SEP-01.  
DR PDB: 1CJF; 26-SEP-01.  
DR PDB: 1CJG; 26-SEP-01.  
DR PDB: 1CJH; 26-SEP-01.  
DR PDB: 1CJI; 26-SEP-01.  
DR PDB: 1CJL; 26-SEP-01.  
DR PDB: 1CJM; 26-SEP-01.  
DR PDB: 1CJN; 26-SEP-01.  
DR PDB: 1CJO; 26-SEP-01.  
DR PDB: 1CJP; 26-SEP-01.  
DR PDB: 1CJQ; 26-SEP-01.  
DR PDB: 1CJR; 26-SEP-01.  
DR PDB: 1CJS; 26-SEP-01.  
DR PDB: 1CJT; 26-SEP-01.  
DR PDB: 1CJU; 26-SEP-01.  
DR PDB: 1CJV; 26-SEP-01.  
DR PDB: 1CJW; 26-SEP-01.  
DR PDB: 1CJX; 26-SEP-01.  
DR PDB: 1CJY; 26-SEP-01.  
DR PDB: 1CJZ; 26-SEP-01.  
DR PDB: 1CJA; 26-SEP-01.  
DR PDB: 1CJB; 26-SEP-01.  
DR PDB: 1CJC; 26-SEP-01.  
DR PDB: 1CJD; 26-SEP-01.  
DR PDB: 1CJE; 26-SEP-01.  
DR PDB: 1CJF; 26-SEP-01.  
DR PDB: 1CJG; 26-SEP-01.  
DR PDB: 1CJH; 26-SEP-01.  
DR PDB: 1CJI; 26-SEP-01.  
DR PDB: 1CJL; 26-SEP-01.  
DR PDB: 1CJM; 26-SEP-01.  
DR PDB: 1CJN; 26-SEP-01.  
DR PDB: 1CJO; 26-SEP-01.  
DR PDB: 1CJP; 26-SEP-01.  
DR PDB: 1CJQ; 26-SEP-01.  
DR PDB: 1CJR; 26-SEP-01.  
DR PDB: 1CJS; 26-SEP-01.  
DR PDB: 1CJT; 26-SEP-01.  
DR PDB: 1CJU; 26-SEP-01.  
DR PDB: 1CJV; 26-SEP-01.  
DR PDB: 1CJW; 26-



Db 92 POCESFDCGKPKVKKPARVVGCVATPISWPAQVSLRRSR-----EHFCGGTLLSP 147  
QY 62 CWVISATHCFIDYPKKEDYVYVGRSLNSNTQGEKMEVENLILHKDYSADTLAHNDI 121  
Db 148 EWLTAARCLDSILGSPFTVILGAHYEMAREASQVEIPVSRLEPESRA-----DI 199  
QY 122 ALLKRSKGRCAQPSRTIQTICLPSMYNDPQF-----GTSCEITGFGKENSVDLYLPBOL 177  
Db 200 ALLKUSP-----AVITDEVIPACLPB-----PNYVADKTVCVITGWGTQGT--FGVGR 249  
QY 178 KMTVVKLISHRECQPPHYGVSEVYTKVCAADPQWKTDSCQDGGPLVCSLQGRMTLTG 237  
Db 250 KEARLPEVINKVGRNRYEVLNGRVKSTELCAGDLAGTDSQDGGPLVCFEKKYILQG 309  
QY 238 IYVSWGRCALCKPKGVYTVVSHFLPMI 264  
Db 310 VTSWGLGRCARPKGVYTVVSHFLPMI 336  
RESULT 36  
TRY2\_BOVIN STANDARD; PRT; 247 AA.  
AC Q29453; 1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Trypsin, anionic precursor (EC 3.4.21.4).  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Holstein-Friesian; TISSUE=Pancreas;  
RX MEDLINE=91065383; PubMed=1701147;  
RA le Hueron I., Wicker C., Guilleau P., Touillet R., Puigserver A.;  
RT "Isolation and nucleotide sequence of cDNA clone for bovine  
pancreatic anionic trypsinogen. Structural identity within the  
trypsin family.";  
RL Eur. J. Biochem. 193:767-773(1990).  
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.  
CC -!- COFACTOR: Binds 1 calcium ion per subunit (By similarity).  
CC -!- SUBCELLULAR LOCATION: Extracellular.  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; X54703; CAA38513.1; -;  
DR PIR; S13813; S13813.  
DR HSSP; P00763; IDPO.  
DR MEROPS; S01.258; -;  
DR InterPro; IPR009003; Cys Ser trypsin.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS50240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolyase; Serine protease; Digestion; Pancreas; Zymogen;  
KW Calcium-binding; signal.  
FT SIGNAL 1 15  
FT PROPEP 16 23  
FT CHAIN 24 247  
ACTIVATION PEPTIDE.  
TRYPSIN, ANIONIC.

FT ACT SITE 63 63 CHARGE RELAY SYSTEM.  
FT METAL 75 75 CALCIUM (BY SIMILARITY).  
FT METAL 77 77 CALCIUM (VIA CARBONYL OXYGEN)  
(BY SIMILARITY).  
FT METAL 80 80 CALCIUM (VIA CARBONYL OXYGEN)  
(BY SIMILARITY).  
FT METAL 85 85 CALCIUM (BY SIMILARITY).  
FT ACT SITE 107 107 CHARGE RELAY SYSTEM.  
FT DISULFID 200 200 CHARGE RELAY SYSTEM.  
FT DISULFID 30 160 BY SIMILARITY.  
FT DISULFID 48 64 BY SIMILARITY.  
FT DISULFID 132 233 BY SIMILARITY.  
FT DISULFID 139 206 BY SIMILARITY.  
FT DISULFID 171 185 BY SIMILARITY.  
FT DISULFID 196 220 BY SIMILARITY.  
FT SITE 194 194 REQUIRED FOR SPECIFICITY.  
SQ SEQUENCE 247 AA; 26289 MW; 50A070495A7731DB CRC64;  
Query Match 25.9%; Score 390.5; DB 1; Length 247;  
Best Local Similarity 38.1%; Pred. No. 4e-28;  
Matches 93; Conservative 41; Mismatches 83; Indels 27; Gaps 9  
QY 23 KIIGGEFTTIENQ-PWFAALYRRHRGGSVTVVCGSLISPCWVISATHCFIDYPKKEDI 81  
Db 23 KIVGG-YTCAENSVPYQVSLNAGY-----HFCGSSLINDQWVVSAAHCY-----QVHIQ 70  
QY 82 VYLGSRSLNSNTQGEKMEVENLILHKDYSADTLAHNDIALILKRSKGRCAQPSRTIQ 141  
Db 71 VRLGEYNDIVLEGGEQFIDASKIRHPKYSSTWL--DNDILLIKLSTP-----AVINARVS 124  
QY 142 TICLPSMYNDPQFGTSCETGFGKENSVDLYPEQLKMTVVKLISHRECQPPHYGVSEVT 201  
Db 125 TLLLPSC--ASAGTECLISGWNLTSSGVNYPDLQLCLVAPLLSHADCEAS--YPOQIT 180  
QY 202 TKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGTVSWGRCALCKPKGVYTVVSHFL 261  
Db 181 NNMICAGFLEGKDKSCQDGGPVACNQ-----LQGVSWGVCQAQKGVYTVVKVCNV 236  
QY 262 PWIR 265  
Db 237 DWIQ 240  
RESULT 37  
CTR2\_CANFA STANDARD; PRT; 263 AA.  
AC P04813.  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Chymotrypsinogen 2 precursor (EC 3.4.21.1).  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=84170253; PubMed=6584866;  
RA Pinsky S.D., Laforge K.S., Luc V., Scheele G.;  
RT "Identification of cDNA clones encoding secretory isoenzyme forms:  
sequence determination of canine pancreatic prechymotrypsinogen 2  
mRNA.";  
RL Proc. Natl. Acad. Sci. U.S.A. 80:7486-7490(1983).  
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-|-Xaa, Trp-|-Xaa,  
Phe-|-Xaa, Leu-|-Xaa.  
CC -!- SUBCELLULAR LOCATION: Extracellular.  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----



```
CC entities requires a license agreement (See http://www.iesib.ch/announce/
CC or send an email to license@iesib.ch).
CC -----
DR EMBL; K01173; AAA30841.1; -.
DR PIR; A21195; A21195.
DR HSSP; P00766; 1ACB.
DR MEROPS; S01152; -.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_DOM; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydroxylase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
FT SIGNAL 1 18
FT CHAIN 19 263 CHYMOTRYPSINOGEN 2.
FT CHAIN 19 31 CHYMOTRYPSIN 2, A CHAIN.
FT CHAIN 34 164 CHYMOTRYPSIN 2, B CHAIN.
FT CHAIN 167 263 CHYMOTRYPSIN 2, C CHAIN.
FT ACT_SITE 75 75 CHARGE RELAY SYSTEM.
FT ACT_SITE 120 120 CHARGE RELAY SYSTEM.
FT ACT_SITE 213 213 CHARGE RELAY SYSTEM.
FT DISULFID 19 140 BY SIMILARITY.
FT DISULFID 60 76 BY SIMILARITY.
FT DISULFID 154 219 BY SIMILARITY.
FT DISULFID 186 200 BY SIMILARITY.
FT DISULFID 209 238 BY SIMILARITY.
SQ SEQUENCE 263 AA; 27787 MW; 2A2F449D813B3961 CRC64;

Query Match 25.9%; Score 390; DB 1; Length 263;
Best Local Similarity 33.2%; Pred. No. 4.8e-28;
Matches 86; Conservative 51; Mismatches 100; Indels 22; Gaps 7;

QY 11 FCGGKTLPRP-----KIIGFTTIENQPFALYRRHGSGVTYVCGSLSPCWVIS 66
DB 17 FCGGVPALQVLSGLSRVNGEDAVPGSNPQVSL-----QDSTGFHFCGSLISEDWVVT 72
QY 67 ATHCFIDYPKEDYIVLGRSLNNTGEMKFEVENILHKDYSADTLAHHNDIALKI 126
DB 73 AAHCGV-----RTHQVAGEFQGGDAESIQLVTKAFKPKFNFMTI--NDITLLKL 126
QY 127 RKEGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGPKENSTDYLPQLKMTVVKLIS 186
DB 127 ATP---ARFSKTVSAVCLPQATDDFPAGTLCVTGWLTKHTNANTPKLQQAALPLLS 182
QY 187 HRECQPHYVYGSVETKMLCAADPQWKTDSCOGDGGPLVCSLQGRMTLTGTVSWGRCA 246
DB 193 NAECKA--FWGSKITDLVWCAGAS--GVSGMGDSGGPLVCCQKAGWTLVGLVSGSGTC 239
QY 247 LKDKPGVYTRVGHFLPWIR 265
DB 239 STSTPGVYARVTKLIPVQ 257

RESULT 38
KLK6_HUMAN STANDARD; PRT; 244 AA.
AC Q92876;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Kallikrein 6 precursor (EC 3.4.21.-) (Protease M) (Neurosin) (Zyme)
DE (SP59).
GN KLK6 OR PRSS9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Carnivora; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RX MEDLINE=97053999; PubMed=8898378;
RA Anisowicz A., Sotiropoulou G., Stenman G., Mok S.C., Sager R.;
RT "A novel protease homolog differentially expressed in breast and
RT ovarian cancer.";
RL Mol. Med. 2:624-636(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=97157069; PubMed=9003450;
RA Yamashiro K., Tsuruoka N., Kodama S., Tsujimoto M., Yamamura Y.,
RA Tanaka T., Nakazato H., Yamaguchi N.;
RT "Molecular cloning of a novel trypsin-like serine protease (neurosin)
RT preferentially expressed in brain.";
RL Biochim. Biophys. Acta 1350:11-14(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97460104; PubMed=9312124;
RA Little S.P., Dixon E.P., Norris F., Buckley W., Becker G.W.,
RA Johnson M., Dobbins J.R., Wyrick T., Miller J.R., Mackellar W.,
RA Hepburn D., Corvalan J., McClure D., Liu X., Stephenson D.,
RA Clemens J., Johnstone E.M.;
RT "Zyme, a novel and potentially amyloidogenic enzyme cDNA isolated
RT from Alzheimer's disease brain.";
RL J. Biol. Chem. 272:25135-25142(1997).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=20079158; PubMed=10610719;
RA Yousef G.M., Luo L.Y., Scherer S.W., Sotiropoulou G., Diamandis E.P.;
RT "Molecular characterization of Zyme/protease M/neurosin (PRSS9), a
RT hormonally regulated kallikrein-like serine protease.";
RL Genomics 62:251-259(1999).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=20510030; PubMed=11054574;
RA Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,
RA Moes P., Paepker B., Wang K.;
RT "Sequencing and expression analysis of the serine protease gene
RT cluster located in chromosome 19q13 region.";
RL Gene 257:119-130(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.S., Toshiruki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.;
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- SUBCELLULAR LOCATION: Secreted
CC -1- TISSUE SPECIFICITY: Preferentially expressed in brain. Also found
CC in colon and kidney.
CC -1- SIMILARITY: Belongs to peptidase family S1. Kallikrein subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collabora-
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on it
CC use by non-profit institutions as long as its content is in no way
```

10-OCT-2003 (Rel. 42, Created)  
10-OCT-2003 (Rel. 42, Last sequence update)  
15-MAR-2004 (Rel. 43, Last annotation update)  
Transmembrane protease, serine 3 (EC 3.4.21.-).

TMPSR3.

Mus musculus (Mouse).

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10909;

[1] SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND FUNCTION IN ENAC  
CLEAVAGE.

MEDLINE=2281255; PubMed=12393794;  
Guipponi M., Vuagniaux G., Wattenhofer M., Shibuya K., Vazquez M.,  
Dougherty L., Scamuffa N., Guida E., Okui M., Rossier C., Hancock M.,  
Buchet K., Raymond A., Hummer E., Marzella P.L., Kudoh J.,  
Shimizu N., Scott H.S., Antonarakis S.E., Rossier B.C.:  
"The transmembrane serine protease (TMPRSS3) mutated in deafness  
DFNB8/10 activates the epithelial sodium channel (ENaC) in vitro.";  
Hum Mol Genet. 11:2829-2836(2002).

CC -!- FUNCTION: Probable protease. Seems to be capable of activating  
ENAC

CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Endoplasmic  
reticulum.

CC -!- TISSUE SPECIFICITY: Expressed in the spiral ganglion, the cells  
supporting the organ of Corti and the stria vascularis.

CC -!- PTM: Undergoes autoproteolytic activation.

CC -!- SIMILARITY: Belongs to peptidase family S1.

CC -!- SIMILARITY: Contains 1 LDL-receptor class A domain.

CC -!- SIMILARITY: Contains 1 SRCR domain.

-----

This SWISS-PROT entry is copyright. It is produced through a collaborati  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
the European Bioinformatics Institute. There are no restrictions on it  
use by non-profit institutions as long as its content is in no wa  
modified and this statement is not removed. Usage by and for commercia  
entities requires a license agreement (See http://www.isb-sib.ch/announce  
or send an email to license@isb-sib.ch).

-----

EMBL; AJ429216; CAD22137.1; --  
EMBL; AJ300738; CAC83350.1; --  
HSSP; P00761; IANI.  
MGD; MG1:2155445; Tmpres3.  
InterPro; IPR009003; Cys Ser trypsin.  
InterPro; IPR002172; LDL\_receptor\_A.  
InterPro; IPR001254; Peptidase\_S1\_  
InterPro; IPR001314; Peptidase\_S1A.  
InterPro; IPR001190; Srcr\_receptor.  
Pfam; PF00057; ldl\_recept\_a; 1.  
Pfam; PF00089; trypsin; 1.  
PRINTS; PR00722; CHYMOTRYPSIN.  
SMART; SM00192; LDLa; 1.  
SMART; SM00202; SR; 1.  
SMART; SM00020; TRYD\_SPC; 1.  
PROSITE; PS01209; LDLRA\_1; 1.  
PROSITE; PS00688; LDLRA\_2; 1.  
PROSITE; PS50287; SRCR\_2; 1.  
PROSITE; PS50240; TRYP SIN DOM; 1.  
PROSITE; PS00134; TRYP SIN HIS; 1.  
PROSITE; PS00135; TRYP SIN SPR; 1.  
Hydrolase; Serine protease; Transmembrane; Signal-anchor; Zymogen;  
Endoplasmic reticulum.

KW DOMAIN 1 48  
FT FT 49 69

TRANSNEM 49 69

-----

DOMAIN 70 453  
FT FT EXTRACELLULAR (POTENTIAL).  
DOMAIN 72 108  
FT FT LDL-RECEPTOR CLASS A.  
DOMAIN 104 205  
FT FT SRCR.  
DOMAIN 217 448  
FT FT SERINE PROTEASE.  
ACT\_SITE 257 257  
FT FT CHARGE RELAY SYSTEM (BY SIMILARITY).  
ACT\_SITE 304 304  
FT FT CHARGE RELAY SYSTEM (BY SIMILARITY).  
ACT\_SITE 400 400  
FT FT CHARGE RELAY SYSTEM (BY SIMILARITY).  
SITE 216 217  
FT FT CLEAVAGE (POTENTIAL).

-----

CYTOPLASMIC (POTENTIAL).  
SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
(POTENTIAL).

EXTRACELLULAR (POTENTIAL).  
LDL-RECEPTOR CLASS A.  
SRCR.  
SERINE PROTEASE.  
CHARGE RELAY SYSTEM (BY SIMILARITY).  
CHARGE RELAY SYSTEM (BY SIMILARITY).  
CHARGE RELAY SYSTEM (BY SIMILARITY).  
CLEAVAGE (POTENTIAL).



Q9E987; Q9ER01;  
 16-OCT-2001 (Rel. 40, Created)  
 16-OCT-2001 (Rel. 40, Last sequence update)  
 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Prostatein precursor (EC 3.4.21.-).  
 GN PRS88.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1].  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Adachi M., Kitamura K., Miyoshi T., Tomita K.;  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2].  
 RP SEQUENCE FROM N.A.  
 RA Wang C.;  
 RT Molecular cloning and expression of rat prostasin.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Possesses a trypsin-like cleavage specificity (By similarity).  
 CC -!- SUBUNIT: Heterodimer of two chains, light and heavy, held by a disulfide bond (By similarity).  
 CC -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND. SECRETED AFTER CLEAVAGE OF ITS C-TERMINUS (BY SIMILARITY).  
 CC -!- SIMILARITY: Belongs to peptidase family S1.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AB017638; BAB20281.1; -.  
 DR EMBL; AF202076; AAG32641.1; -.  
 DR HSSP; P00734; 1UVS.  
 DR InterPro; IPR009003; Cys Ser trypsin.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS50240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Hydrolyase; Serine protease; Zymogen; Signal; Glycoprotein;  
 KW Transmembrane.  
 FT SIGNAL 1 29 POTENTIAL.  
 FT PROPEP 30 32 ACTIVATION PEPTIDE (BY SIMILARITY).  
 FT CHAIN 33 44 PROSTASIN LIGHT CHAIN.  
 FT CHAIN 45 322 PROSTASIN HEAVY CHAIN.  
 FT PROPEP 323 342 BY SIMILARITY.  
 FT TRANSMEM 320 340 POTENTIAL.  
 FT DOMAIN 45 286 SERINE PROTEASE.  
 FT DISULFID 37 154 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 70 86 BY SIMILARITY.  
 FT DISULFID 168 244 BY SIMILARITY.  
 FT DISULFID 201 223 BY SIMILARITY.  
 FT ACT\_SITE 85 85 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 134 134 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 238 238 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT CARBOHYD 159 159 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CONFLICT 17 17 I -> V (IN REF. 1).  
 FT CONFLICT 292 292 A -> V (IN REF. 1).  
 SQ SEQUENCE 342 AA; 36843 MW; 5EDIAF05D2113B98 CRC64;  
 Query Match 25.6%; Score 386.5; DB 1; Length 342;  
 Best Local Similarity 36.3%; Pred. No. 1.4e-27;  
 Matches 99; Conservative 29; Mismatches 108; Indels 37; Gaps 9;

QY 13 CGQKTLRPRFKLIGGEFTTIENQWPFAAIYRRHGGSVTY-----VCGGSLISPCWVISAT 68  
 DB 37 CG-AVIQPR--ITGGSAKPGQWPNQV-----SITYGVHVCGLSVNQWVWSA 84  
 QY 69 HCFIDYPKEDYIVYLGSRSLNSNTQGENKFEVENLILHKOYADTLAHNDIALKIRS 128  
 DB 85 HCFPRHSKEEVEVKLGALQDPSFNDIVVHTVAQIISHSSYREE--GSGQDIALIRLS 142  
 QY 129 KEGRCAQPSRTIQTICLPSMYNDPQGTSCETITGFGK-ENSTDYLYPEOLKMTVVKLISH 187  
 DB 143 P-----VTFERYIRPICLPANASFPNGLHCTVTGHWVAPSLSLQTPRPLQOLEVPLISR 198  
 QY 188 REC-----QPHYHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRWTLTG 237  
 DB 199 ETGCLYNINAVPEEPH-----TIQDMLCAGYVKGKDACQDGGSPISCPDGLWYLAG 254  
 QY 238 IYSWGEGCALDKDGPVYTVRSHLPKIRSHKE 270  
 DB 255 IYSWGACGAPRPGVITLTSTYASWIIHVAE 287  
 RESULT 42  
 HEPS RAT  
 ID HEPS RAT STANDARD; PRT; 416 AA.  
 AC Q05511;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Serine protease hepsin (EC 3.4.21.-).  
 GN HPN.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1].  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=93305733; PubMed=8318546;  
 RA Farley D., Raymond F., Nick H.;  
 RT Cloning and sequence analysis of rat hepsin, a cell surface serine proteinase.;  
 RL Biochim. Biophys. Acta 1173:350-352(1993).  
 CC -!- FUNCTION: Plays an essential role in cell growth and maintenance of cell morphology.  
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein.  
 CC -!- SIMILARITY: Belongs to peptidase family S1.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X70900; CAA50256.1; -.  
 DR PIR; S33777; S33777.  
 DR HSSP; P00763; IDPO.  
 DR MEROPS; S01.224; -.  
 DR InterPro; IPR009003; Cys Ser trypsin.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR InterPro; IPR001190; Srcr\_receptor.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00202; SR; 1.  
 DR SMART; SM00202; Tryp\_SPC; 1.  
 DR PROSITE; PS50240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Hydrolyase; Serine protease; Transmembrane; Signal-anchor.  
 FT CHAIN 1 161 SERINE PROTEASE HEPsin, NON-CATALYTIC

```
FT CHAIN 162 416 CHAIN (POTENTIAL).
FT SERINE PROTEASE HEPSIN, CATALYTIC CHAIN
FT (POTENTIAL).
FT CYTOPLASMIC (POTENTIAL).
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT TRANSMEM 17 43 (POTENTIAL).
FT EXTRACELLULAR (POTENTIAL).
FT DOMAIN 44 416
FT DOMAIN 162 416
FT ACT SITE 202 256
FT ACT SITE 256 256
FT ACT SITE 352 352
FT ACT SITE 352 352
FT DISULFID 152 276
FT DISULFID 187 203
FT DISULFID 321 337
FT DISULFID 348 380
FT CARBOHYD 111 111
FT CARBOHYD 416 AA; 44926 MW; ESA9F8FA9550E180 CRO64;
SQ SEQUENCE 25.6%; Score 386.5; DB 1; Length 416;
Best Local Similarity 35.0%; Pred. No. 1.se-27;
Matches 99; Conservative 51; Mismatches 90; Indels 43; Gaps 14;

QY 13 CSQKTLRPRFKIIGGFTTIENQWPAALYRHRGGSVTVYCGSLISPCWVISATHCFI 72
DB 152 CGRRKL-PVDRIVGGQSSLRWPMQVSL--RYDG---THLCGSLSGDWLTAHCF- 204
QY 73 DYPKKEDYI---VYLGRSLNSNTGEMKFEVENLILH-----KQYSADTLAHHNDIA 122
DB 205 --PERNVLSRWFVAG--AVARTSPHVLGVQAVHYGGVLPFRDPTID--ENSNDIA 258
QY 123 LKIRSEKRCQAQPSRTIITICLPSMYNDPQFGTSCETGFGKENSTYLYPEQ---LKM 179
DB 259 LVHLSSS---LPLTEYIQVCLPAAGQALVDGKVCVTGWTGNTQ---FYQQAVWLQE 310
QY 180 TVVKYL-SHRECPQPHYVYGGSEVTMLCAADPQWKTDSCQDGGGPIVC-----SLOGMTL 235
DB 311 ARVPIISNEVCNSPDYFGNQIKPMFCAGYPEGDAGCQDGGGHHFVCDRI-SGTSRWL 370
QY 236 TGIVSGRGKALCKDKPGVTVTRVSHPLPW----IRSHTKEENGL 274
DB 371 CGIVSGTGCALARKPGVTVKVIDPREWIFQAIKTHS-EATGM 412

RESULT 43
KAL_RAT STANDARD; PRT; 638 AA.
ID_KAL_RAT
AC P14272;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Plasma kallikrein precursor (EC 3.4.21.34) (Plasma prekallikrein)
DE (Kininogenin) (Fletcher factor).
GN KUKB1 OR PK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91129236; PubMed=1993180;
RA Beaubien G., Rosinski-Chupin I., Mattei M.-G., Mbikay M., Chretien M.,
RA Seidah N.G.;
FT "Gene structure and chromosomal localization of plasma kallikrein.";
RL Biochemistry 30:1628-1635(1991).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=90091743; PubMed=2598771;
RA Seidah N.G., Ladenheim R., Mbikay M., Hamelin J., Lutfalla G.,
RA Rougeon F., Lazure C., Chretien M.;
FT "The cDNA structure of rat plasma kallikrein.";
RL DNA 8:563-574(1989).
CC -!- FUNCTION: The enzyme cleaves Lys-Arg and Arg-Ser bonds. It
activates, in a reciprocal reaction, factor XII after its binding
```

```
CC to a negatively charged surface. It also releases bradykinin from
CC HMW kininogen and may also play a role in the renin-angiotensin
CC system by converting prorenin into renin.
CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Xaa and Lys-|-Xaa
CC bonds, including Lys-|-Arg and Arg-|-Ser bonds in (human)
CC kininogen to release bradykinin.
CC -!- SUBUNIT: The zymogen is activated by factor XIIa, which cleaves
CC the molecule into a light chain, which contains the active site,
CC and a heavy chain, which associates with HMW kininogen. These
CC chains are linked by one or more disulfide bonds.
CC -!- SIMILARITY: Belongs to peptidase family S1. Plasma kallikrein
CC subfamily.
CC -!- SIMILARITY: Contains 4 apple domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboratio
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on it
CC use by non-profit institutions as long as its content is in no wa
CC modified and this statement is not removed. Usage by and for commercia
CC entities requires a license agreement (see http://www.isb-sib.ch/announce
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M62357; AAA74563.1; --
CC EMBL; M62358; AAA74563.1; JOINED.
CC EMBL; M62346; AAA74563.1; JOINED.
CC EMBL; M62347; AAA74563.1; JOINED.
CC EMBL; M62349; AAA74563.1; JOINED.
CC EMBL; M62350; AAA74563.1; JOINED.
CC EMBL; M62351; AAA74563.1; JOINED.
CC EMBL; M62352; AAA74563.1; JOINED.
CC EMBL; M62353; AAA74563.1; JOINED.
CC EMBL; M62354; AAA74563.1; JOINED.
CC EMBL; M62355; AAA74563.1; JOINED.
CC EMBL; M62356; AAA74563.1; JOINED.
CC EMBL; M30282; AAA41463.1; --
CC EMBL; M58590; AAA42069.1; --
CC PIR; A39180; KQTPPL.
CC HSSP; P00750; IRTF.
CC MEROPS; S01.212; --
CC InterPro; IPR000177; Apple.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR003014; PAN.
CC InterPro; IPR001254; Peptidase_S1.
CC InterPro; IPR001314; Peptidase_S1A.
CC Pfam; PF00024; PAN; 4.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00005; APPLEDOMAIN.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC SMART; SM00223; APPLE; 4.
CC SMART; SM00020; Tryp_SPC; 1.
CC PROSITE; PS00495; APPLE; 4.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC Hydrolase; Serine protease; Glycoprotein; Plasma; Zymogen; Signal;
KW Fibrinolysis; Blood coagulation; Inflammatory response; Liver;
KW Repeat.
FT SIGNAL 1 19
FT CHAIN 20 390 PLASMA KALLIKREIN HEAVY CHAIN.
FT CHAIN 391 638 PLASMA KALLIKREIN LIGHT CHAIN.
FT DOMAIN 20 105 APPLE 1.
FT DOMAIN 110 195 APPLE 2.
FT DOMAIN 200 285 APPLE 3.
FT DOMAIN 291 376 APPLE 4.
FT DOMAIN 389 621 SERINE PROTEASE.
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 308 308 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 453 453 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 494 494 N-LINKED (GLCNAC. . .) (PROBABLE).
FT ACT_SITE 434 434 CHARGE RELAY SYSTEM.
FT ACT_SITE 483 483 CHARGE RELAY SYSTEM.
FT ACT_SITE 578 578 CHARGE RELAY SYSTEM.
```

DT	DISULFID	21	104	BY SIMILARITY.
FT	DISULFID	47	77	BY SIMILARITY.
FT	DISULFID	51	57	BY SIMILARITY.
FT	DISULFID	111	194	BY SIMILARITY.
FT	DISULFID	137	166	BY SIMILARITY.
FT	DISULFID	141	147	BY SIMILARITY.
FT	DISULFID	201	284	BY SIMILARITY.
FT	DISULFID	227	255	BY SIMILARITY.
FT	DISULFID	231	237	BY SIMILARITY.
FT	DISULFID	292	375	BY SIMILARITY.
FT	DISULFID	318	347	BY SIMILARITY.
FT	DISULFID	322	328	BY SIMILARITY.
FT	DISULFID	340	345	BY SIMILARITY.
FT	DISULFID	383	503	BY SIMILARITY.
FT	DISULFID	419	435	BY SIMILARITY.
FT	DISULFID	517	584	BY SIMILARITY.
FT	DISULFID	548	563	BY SIMILARITY.
FT	DISULFID	574	602	BY SIMILARITY.
SEQ	SEQUENCE	638 AA;	71273 MW;	454BEB276CASP88 CRC64;

Query Match 25.5%; Score 385; DB 1; Length 638;  
Best Local Similarity 32.3%; Pred. No. 4.1e-27;  
Matches 85; Conservative 56; Mismatches 106; Indels 16; Gaps 3;

QY	17	TLRPFRIIGBEFTTIENQWFAAIYRRHRGGSVTYCGGSLISPCWVISATHCFIDYPK	76
Ddb	384	TTKINARIVGTSNGIGENFWQVSL--QVKLVSONHMKCGSIGIGWILTAHCFDGPY	441
QY	77	KEDYIVYGLGRSLNSNTQGEKMEKFEVENLILHKDYSADTLAHRNDIALLKIRSKGRCAQP	136
Ddb	442	PDVMRIYGGIILNISEITNKTFFPSIKELIHHQRYKMGESY--DIALIKLQTP----	495
QY	137	SRTIQTICLPSMYNDPQFGSCITGFG--KENSNDYLYPQLKNTVVYKLISHRECQP	193
Ddb	496	TEFOKPCILPSKADNTIYINCWTGWGYTKERGETQNI----LQKATIPLVNECOQ-	550
QY	194	HYGSEVTTKMLCAADPQWKTDSCQSGGGLVCSLQGRMTUTGIVSMRGCGALKDKPGV	253
Ddb	551	KYRDVITKMICAGYKEGGIDACKGDSGGPLVCKHSGRWLVGITSWGEGCARKEQPGV	610
QY	254	TRYSHPLPWRSHKTEENGAL	276
Ddb	611	YTKVAEYIDWILEKIQSKERAL	633

RESULT 44  
TRYG MOUSE  
ID TRYG MOUSE STANDARD; PRT; 311 AA.  
AC O9QUL7;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Tryptase gamma precursor (EC 3.4.21.-) (Transmembrane tryptase).  
GN TPSS1 OR TWI.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
RN NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/SV, and BALB/c;  
RX MEDLINE=99452974; PubMed=10521469;  
RA Wong G.W., Tang Y., Feyfant E., Sali A., Li L., Li Y., Huang C.,  
RA Friend D.S., Krilis S.A., Stevens R.L.;  
RT "Identification of a new member of the tryptase family of mouse and  
RT human mast cell proteases which possesses a novel COOH-terminal  
RT hydrophobic extension."  
RL J. Biol. Chem. 274:30784-30793 (1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FVB/N; TISSUE=Colon;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,





DB	127	ATP----	AQPSQSVSAVCLPNVDDPPGTVCATTCGKTKYNALKTPEKLAQALPIVS	187
QY	187	HRECOQPHYGYSEVTTKXLCALDPQKWKTDSCQDQSGPLVCSLQGNWTLTGIVSWGRCA	246	
DB	183	EADCKKS--WGSKITDVTMT	CAGAS--GVSSCMGDSGGLPVCCQKDGVTMTLAGIVSWGSGVC	238
QY	247	LKDQPGVYTRVSHFLPMIR	265	
DB	239	STSTPANSYSRVTALPMFVQ	257	
RESULT 47				
ID	PS58	MOUSE	STANDARD;	PRT; 342 AA.
AC	Q9E9D1			
DT	16-OCT-2001	(Rel. 40, Created)		
DT	16-OCT-2001	(Rel. 40, Last sequence update)		
DE	15-MAR-2004	(Rel. 43, Last annotation update)		
DE	Prostasin precursor (EC 3.4.21.-)	(Channel activating protease 1).		
GS	PRSS8	OR CAP1.		
OS	Mus musculus	(Mouse).		
OC	Eukaryota;	Metazoa;	Chordata;	Craniata; Vertebrata; Euteleostomi;
OC	Mammalia;	Eutheria;	Rodentia;	Sciurognathi; Muridae; Murinae; Mus.
NCBI	Taxid=10090;			
[1]				
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=2035202;	PubMed=10770960;		
RX	Vuagniaux G, Vallet V, Jaeger N.F., Pfister C., Bens M., Farman N.,			
RA	Coutois-Couty N, Vandewalle A, Rossier B.C, Hummler E.			
RT	Activation of the amiloride-sensitive epithelial sodium channel by			
RT	the serine protease mCAP1 expressed in a mouse cortical collecting			
RT	duct cell line."			
RL	J. Am. Soc. Nephrol. 11:829-834 (2000).			
CC	-I- FUNCTION: Possesses a trypsin-like cleavage specificity (By			
CC	similarity). Activates amiloride-sensitive sodium channels.			
CC	-I- SUBUNIT: Heterodimer of two chains, light and heavy, held by a			
CC	disulfide bond (By similarity).			
CC	-I- SUBCELLULAR LOCATION: MEMBRANE-BOUND. SECRETED AFTER CLEAVAGE OF			
CC	ITS C-TERMINUS (BY SIMILARITY).			
CC	-I- SIMILARITY: Belongs to peptidase family S1.			
CC	-I- CAUTION: Ref.1 sequence differs from that shown due to a			
CC	frameshift in position 339.			
CC				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	at the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announcement">http://www.isb-sib.ch/announcement</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC				
DR	EMBL; AF188613;	AG17054.1;	ALT_FRAME.	
DR	HSSP; P00734;	LUVS.		
DR	MEROPS; S01.158;	-		
DR	MGD; MG1:1923810;	Prss8.		
DR	InterPro; IPR009003;	Cys Ser trypsin.		
DR	InterPro; IPR001254;	Peptidase S1.		
DR	InterPro; IPR001314;	Peptidase_S1A.		
DR	Pfam; PF00089;	Crypsin; 1.		
DR	PRINTS; PR00722;	CHYMOTRYPSIN.		
DR	SMART; SM00020;	Tryp_SPC; 1.		
DR	PROSITE; PS50240;	TRYPSIN_DOM; 1.		
DR	PROSITE; PS00134;	TRYPSIN_HIS; 1.		
DR	PROSITE; PS00135;	TRYPSIN_SER; 1.		
KW	Hydrolase; serine protease; Zymogen; Signal; Glycoprotein;			
KW	Transmembrane.			
FT	SIGNAL	1	29	POTENTIAL.
FT	PROPEP	30	32	ACTIVATION PEPTIDE (BY SIMILARITY).
FT	CHAIN	33	44	PROSTATIN LIGHT CHAIN.
FT	CHAIN	45	322	PROSTATIN HEAVY CHAIN.
FT	PROPEP	323	342	BY SIMILARITY.
FT	TRANSMEM	320	340	POTENTIAL.
FT	DOMAIN	45	286	SERINE PROTEASE.



FT DISULFID 37 154 INTERCHAIN (BY SIMILARITY).  
FT DISULFID 70 86 BY SIMILARITY.  
FT DISULFID 168 244 BY SIMILARITY.  
FT DISULFID 201 223 BY SIMILARITY.  
FT DISULFID 234 262 BY SIMILARITY.  
FT ACT\_SITE 85 85 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 134 134 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 238 238 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 159 159 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 342 AA; 36729 MW; 0620DB88ED187DOF CRC64;  
  
Query Match 25.2%; Score 380.5; DB 1; Length 342;  
Best Local Similarity 35.5%; Pred. No. 4.8e-27;  
Matches 97; Conservative 30; Mismatches 109; Indels 37; Gaps 9;  
  
QY 13 CGOKTRPRFKIIGGFTTIENQPFALYRRHGGSVTY-----VCGSLISPCWVIGAT 68  
DB 37 CG-AVQPR-ITGGGSAKPGQWPQV-----SITYDGNHVCGGSLVSNKVVSA 84  
  
QY 69 HCFIDYPKEDYIVYGRSLNSNTGEMKFEVENILHKDYSADTLAHNDLALKIRSG 128  
DB 85 HCFPREHREAYEVKLGALHQLDSYSDNTVHTVAQIITHSSYREE--GSGQDIAFIRLSS 142  
  
QY 129 KEGRCAQPRRTIOTICLPNMYNDPQFGTSCETITGPK-ENSTDYLYPEOLKMTVVKLISH 187  
DB 143 P-----VTFSRVIRPCLPANASFPNGLHCTVTGNGHVAPSVSLQTPRFLQLEVLISR 198  
  
QY 188 REC-----QQHYHGVSEVTVKALCAADPQWKTSCQDSGSGPLVCSLQGRMTLTG 237  
DB 199 ETCSCLYINNAVPEEPH-----TIQDMLCAGYVKGKDACQDSDGSGPLSCPMEGIWYLAG 254  
  
QY 236 IVSWGRGCALKDKPGVYTVRSHLPWIRSHKTE 270  
DB 255 IVSWGDCAGAPNPGVTVTLTSTYASWIIHHVAE 287  
  
RESULT 48  
EL2\_PIG STANDARD; PRT; 269 AA.  
AC P08419;  
DT 01-AUG-1998 (Rel. 08, Created)  
DT 01-AUG-1998 (Rel. 08, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Elastase 2 precursor (EC 3.4.21.71).  
GN ELA2.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87217962; PubMed=3646943;  
RA Kawashima I., Tani T., Shimoda K., Takiguchi Y.;  
RT "Characterization of pancreatic elastase II cDNAs: two elastase II  
RT mRNAs are expressed in human pancreas.";  
RL DNA 6:163-172(1987).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88198076; PubMed=2834346;  
RA Shirasu Y., Yoshida H., Matsuki S., Takemura K., Ikeda N.,  
RA Shimada Y., Ozawa T., Mikayama T., Iijima H., Ishida A., Sato Y.,  
RA Tanai Y., Tanaka J., Ikenaga H.;  
RT "Molecular cloning and expression in Escherichia coli of a cDNA  
RT encoding human pancreatic elastase 2.";  
RL J. Biochem. 102:1555-1563(1987).  
CC -|- FUNCTION: Acts upon elastin.  
CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Leu|-Xaa, Met|-Xaa  
CC and Phe|-Xaa. Hydrolyzes elastin.  
CC -|- SUBCELLULAR LOCATION: Secreted.  
CC -|- TISSUE SPECIFICITY: Pancreas.  
CC -|- SIMILARITY: Belongs to peptidase family S1. Elastase subfamily.  
-----

This SWISS-PROT entry is copyright. It is produced through a collaboratio  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
the European Bioinformatics Institute. There are no restrictions on it  
use by non-profit institutions as long as its content is in no wa  
modified and this statement is not removed. Usage by and for commercia  
entities requires a license agreement (See http://www.isb-sib.ch/announce  
or send an email to license@sib-sib.ch).  
-----  
DR EMBL; M16551; AAA31027.1; -;  
DR EMBL; D00237; BAA00166.1; -;  
DR PIR; A26823; A26823.  
DR PDB; 1BRU; 25-AUG-99.  
DR MEROPS; S01.155; -;  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; Tryp\_Spc; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease; Zymogen; Signal; 3D-structure.  
FT SIGNAL 1 16  
FT PROPEP 17 28 ACTIVATION PEPTIDE.  
FT CHAIN 29 269 ELASTASE 2.  
FT ACT\_SITE 73 73 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 121 121 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 216 216 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT DISULFID 58 74 BY SIMILARITY.  
FT DISULFID 155 222 BY SIMILARITY.  
FT DISULFID 212 243 BY SIMILARITY.  
FT CONFLICT 10 10 L -> S (IN REF. 2).  
FT CONFLICT 118 118 N -> K (IN REF. 2).  
FT CONFLICT 132 132 S -> Y (IN REF. 2).  
FT CONFLICT 172 172 I -> V (IN REF. 2).  
FT CONFLICT 202 202 C -> V (IN REF. 2).  
SQ SEQUENCE 269 AA; 28699 MW; BAC6FE9AF4DDE56 CRC64;  
  
Query Match 25.2%; Score 380; DB 1; Length 269;  
Best Local Similarity 34.1%; Pred. No. 4e-27;  
Matches 88; Conservative 50; Mismatches 106; Indels 14; Gaps 8  
  
QY 13 CGOKTRPRF-KITGGFTTIENQPFALYRRHGGSVTYVCGSLISPCWVIGAT 71  
DB 17 CGLPANLPQLPVVGEDARFNSWPQVSL-QYDSSGQWRTCGTILVDQSWLTAARCI 75  
  
QY 72 IDYPKEDYIVYGRSLNSNTGEMKFEVENILHKDYSADTLAHNDLALKIRSG 131  
DB 76 ---SSSRTRYVVLGRHSLSTNEPGLAVKSLVHQQDWSNQLSNGNDLALKLSP-- 130  
  
QY 132 RCAQPSRTIOTICLPNMYNDPQFGTSCETITGPKENSTDYLYPEOLKMTVVKLISHRECQ 191  
DB 131 ---VSLTDKIQGLCLPAAGTILPNNVYCVYTGWR-LQINGASPDILQOGLLVVDYATCS 187  
  
QY 192 QPHYHGVSEVTVKALCAADPQWKTSCQDSGSGPLVCSLQGRMTLTGIVSWGR--GCALK 248  
DB 188 KPGWVGSTVKTNTMTCAGG-DGIISSCNGDSGGLNCGQANGQWQVHGIVSPGSLGNY 246  
  
QY 249 DKPGYTVRVSHFLPWIRS 266  
DB 247 HKPSVFTVSNYIDWINS 264  
  
RESULT 49  
TRY3 HUMAN STANDARD; PRT; 304 AA.  
ID TRY3 HUMAN  
AC P35030; P35951; Q15665; Q9UQV3;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Trypsin III precursor (EC 3.4.21.4) (Brain trypsinogen)



RA Larsen F., Solheim J., Kristensen T., Kolsto A.B., Prydz H.;  
RT "A tight cluster of five unrelated human genes on chromosome  
EL 16q22.1"; Hum. Mol. Genet. 2:1589-1595(1993).  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X71874; CAA50710.1; -;  
CC EMBL; X71877; CAA50711.1; -;  
CC PIR; I38136; I38136.  
CC HSSP; P00763; LDPO.  
CC MEROPS; S01.256; -;  
CC Genew; HGNC:2524; CTRL.  
CC  
CC GO; GO:0005615; C:extracellular space; TAS.  
CC GO; GO:0007586; P:digestion; TAS.  
CC GO; GO:0006508; P:proteolysis and peptidolysis; TAS.  
CC InterPro; IPR009003; Cys Ser trypsin.  
CC InterPro; IPR001254; Peptidase\_S1.  
CC InterPro; IPR001314; Peptidase\_S1A.  
CC Pfam; PF00089; trypsin; 1.  
CC PRINTS; PR00722; CHYMOTRYPSIN.  
CC SMART; SM0020; Tryp\_SPC; 1.  
CC PROSITE; PS0240; TRYPSIN\_DOM; 1.  
CC PROSITE; PS00134; TRYPSIN\_HIS; 1.  
CC PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydroxylase; Serine protease; Glycoprotein; Zymogen; Signal.  
FT SIGNAL 1 18  
FT PROPEP 19 33  
FT CHAIN 34 264  
FT ACT\_SITE 75 75  
FT ACT\_SITE 121 121  
FT ACT\_SITE 214 214  
FT CARCHYD 114 114  
FT DISULFID 19 141  
FT DISULFID 60 76  
FT DISULFID 155 220  
FT DISULFID 187 201  
FT DISULFID 210 239  
SQ SEQUENCE 264 AA; 28002 MW; 3F629F02FA6DDFB4 CRC64;  
Query Match 25.1%; Score 378.5; DB 1; Length 264;  
Best Local Similarity 35.0%; Pred. No. 5.3e-27;  
Matches 90; Conservative 43; Mismatches 101; Indels 23; Gaps 10;  
QY 13 CGQKTLRPRF---XIIIGGETTIENQWFAAIVRRHGGSVTVVCGGSLISPCWVISAT 68  
DB 19 CGIPAKPALSPSRIVNGENAVLGSPWQVSL-----QDSGFFHFCGSLISQSWVTAA 74  
QY 69 HCFIDYPKEDYIVYLGRSLNSNTQGMKFEVENLIHKYSDATLAHNDIALKTRIS 128  
DB 75 HCNVS-PGR--HFVVLGEYDRSSNAEPLQLVSVSRAITHPSWNSTTM--NNDVTLLKLS 129  
QY 129 KEGRCAPQRTTQTCLPSMYNDPOFGTSCETFGKENSITDLYPEOLKMTVVKLISHR 188  
DB 130 P----AQYTRISPCVCLASNEALTEGLTCVTIGMGRUGVGNVTPAHLQVALPLVIVN 185  
QY 189 EQQPHYGVSEVTKMLCAADPQWKTDSCQGSQGGPLVCSLQGRMTLTGIVSWG-RGCAL 247  
DB 186 QCRQ--YWGSSITDSMICAGGA--GASSCQGSQGGPLVCSLQGRMTLTGIVSWGKNCNV 241  
QY 248 KKKPGVYTRVSHLPWI 264  
DB 242 R-APAVYTRVSKFTWI 257

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2004, 14:44:05 ; Search time 45.7178 Seconds

(without alignments)  
1904.795 Million cell updates/sec

Title: US-09-880-503-5

Perfect score: 1508

Sequence: 1 KPSSPPEELFKCQKTLRP.....VSHFLPWIRSHTEKNGAL 276

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 75 summaries

Database :

SPTREMBL\_25.\*

1: sp\_archaea.\*

2: sp\_bacteria.\*

3: sp\_fungi.\*

4: sp\_human.\*

5: sp\_invertebrate.\*

6: sp\_mammal.\*

7: sp\_mhc.\*

8: sp\_organelle.\*

9: sp\_phage.\*

10: sp\_plant.\*

11: sp\_rendent.\*

12: sp\_virus.\*

13: sp\_vertebrate.\*

14: sp\_unclassified.\*

15: sp\_rvirus.\*

16: sp\_bacteriaph.\*

17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1321	87.6	433	6 Q8MIL0	Q8milo cryptolagus
2	1316	87.3	433	6 Q8MHI7	Q8mhi7 cryptolagus
3	670	44.4	214	6 Q9XT70	Q9xt70 cryptolagus
4	585	38.8	128	6 Q97587	Q97587 cryptolagus
5	583	38.7	516	4 Q9BU99	Q9bu99 homo sapien
6	583	38.7	562	4 Q85YK8	Q85yk8 homo sapien
7	573	38.0	395	4 Q9BZW1	Q9bzw1 homo sapien
8	553	36.7	562	6 Q8SQ23	Q8sq23 sus scrofa
9	545.5	36.2	564	6 Q8MKG1	Q8mkg1 cryptolagus
10	497.5	33.0	653	11 Q8VCS4	Q8vcs4 mus musculus
11	495	32.8	103	6 Q9SM89	Q9sm89 equus caball
12	495	32.8	300	4 Q96EF3	Q96ef3 homo sapien
13	492	32.6	540	13 Q800V7	Q800v7 meleagris g
14	492	32.6	615	4 Q81ZZ5	Q81zz5 homo sapien
15	467	31.0	616	6 Q97507	Q97507 sus scrofa
16	463.5	30.7	868	5 Q9YIV3	Q9yiv3 polyandroca

# ALIGNMENTS

## RESULT 1

ID	Q8MIL0	PRELIMINARY;	PRT;	433 AA.
ID	Q8MIL0			
AC	Q8MIL0			
DT	01-OCT-2002 (TEMBUREL. 22, Last sequence update)			
DT	01-OCT-2002 (TEMBUREL. 22, Last sequence update)			
DT	01-OCT-2003 (TEMBUREL. 25, Last annotation update)			
DE	Urokinase-type plasminogen activator.			
GN	PLAU.			

17	451.5	29.9	597	11	Q35727	Q35727 mus musculus
18	451.5	29.9	609	11	Q80YC5	Q80yc5 mus musculus
19	442.5	29.3	327	4	Q8N171	Q8n171 homo sapien
20	435.5	28.9	284	4	Q8NF86	Q8nf86 homo sapien
21	426.5	28.3	267	5	Q9BK47	Q9bk47 luidia foli
22	425.5	28.2	505	5	Q966V4	Q966v4 halocynthia
23	425	28.2	517	11	Q8K0D2	Q8k0d2 mus musculus
24	421.5	28.0	761	11	Q89JC8	Q89jc8 rattus norv
25	420	27.9	471	11	Q8CF80	Q8cfe0 mus musculus
26	418.5	27.8	277	11	Q80WM7	Q80wm7 mus musculus
27	418	27.7	537	4	Q8BYE1	Q8bye1 homo sapien
28	418	27.7	558	4	Q86YMA	Q86yma homo sapien
29	417	27.7	560	4	Q14520	Q14520 homo sapien
30	414	27.5	581	4	Q8BYE2	Q8bye2 homo sapien
31	411.5	27.3	276	11	Q8CGR6	Q8cgr6 mus musculus
32	411	27.3	538	11	Q8R0P5	Q8r0p5 mus musculus
33	402.5	26.7	371	11	Q8CJ16	Q8cjl6 rattus norv
34	402.5	26.7	445	11	Q8CJ17	Q8cjl7 rattus norv
35	396.5	26.3	455	11	Q8CDR0	Q8cdr0 mus musculus
36	396	26.3	1059	4	Q72411	Q72411 homo sapien
37	394.5	26.2	277	5	Q96899	Q96899 scolopendra
38	391.5	26.0	261	4	Q725F4	Q725f4 homo sapien
39	390	25.9	855	4	Q72410	Q72410 homo sapien
40	389.5	25.8	453	11	Q812A6	Q812a6 mus musculus
41	384.5	25.5	339	11	Q99L44	Q99l44 mus musculus
42	384	25.5	340	11	Q8BJV6	Q8bjv6 mus musculus
43	383	25.4	263	11	Q9DCR6	Q9dc86 mus musculus
44	382	25.3	263	11	Q9CR35	Q9cr35 mus musculus
45	382	25.3	624	11	Q9DAT3	Q9dat3 mus musculus
46	381	25.3	249	11	Q9QYNA	Q9qyn4 mus musculus
47	381	25.3	276	11	Q9QYN3	Q9qyn3 m hippostas
48	381	25.3	767	13	Q9DGR2	Q9dgr2 xenopus lae
49	379.5	25.2	251	4	Q8NZU3	Q8nzu3 homo sapien
50	378.5	25.1	269	4	Q8IUW0	Q8iuw0 homo sapien
51	377.5	25.0	429	13	Q8AVB0	Q8avb0 brachydanio
52	377.5	25.0	681	13	Q72T70	Q72t70 lampetra ja
53	377	25.0	321	4	Q96RZ8	Q96rz8 homo sapien
54	377	25.0	707	13	Q8CGV0	Q8cgv0 cyprinus ca
55	376	24.9	263	11	Q9D8X8	Q9d8x8 mus musculus
56	376	24.9	624	11	Q91Y47	Q91y47 mus musculus
57	375	24.9	812	11	Q9R0W3	Q9r0w3 rattus norv
58	374	24.8	371	5	Q8MRY3	Q8mry3 drosophila
59	374	24.8	1374	5	Q9VSU0	Q9vsu0 drosophila
60	374	24.8	1449	5	Q9ULI2	Q9uli2 drosophila
61	374	24.8	1450	5	Q8ICB8	Q8icb8 drosophila
62	374	24.8	1462	5	Q9ULI3	Q9uli3 drosophila
63	374	24.8	2382	5	Q9B1I9	Q9b1i9 drosophila
64	374	24.8	2409	5	Q960G6	Q960g6 drosophila
65	374	24.8	2786	5	Q9VSU2	Q9vsu2 drosophila
66	373.5	24.8	264	11	Q9D7P8	Q9d7p8 mus musculus
67	373.5	24.8	264	11	Q9D960	Q9d960 mus musculus
68	373.5	24.8	264	11	Q9ER05	Q9er05 mus musculus
69	373	24.7	315	5	Q8IRR3	Q8irr3 drosophila
70	372.5	24.7	643	6	Q97506	Q97506 sus scrofa
71	372.5	24.7	247	11	Q9CPN9	Q9cpn9 mus musculus
72	372.5	24.7	248	13	Q7SZT1	Q7szt1 xenopus lae
73	372.5	24.7	572	11	Q8BIK6	Q8bik6 mus musculus
74	372.5	24.7	1322	5	Q9NAT0	Q9nat0 anopheles g
75	371.5	24.6	1322	5	Q9NJS5	Q9nj95 anopheles g

OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22155945; PubMed=12149463;  
RA Falkenberg M., Tom C., DeYoung M.B., Wen S., Linnemann R.,  
RA Dichek D.A.;  
RT "Increased expression of urokinase during atherosclerotic lesion  
RT accelerates arterial constriction and lumen loss, and  
RL Proc. Natl. Acad. Sci. U.S.A. 99:10665-10670(2002).  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  
DR EMBL; AY122285; AAM83187.1; -.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0016301; F:kinase activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:000508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR009003; Cys Ser trypsin.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR001254; Peptidase S1.  
DR InterPro; IPR001314; Peptidase S1A.  
DR InterPro; IPR008293; Pept\_S1A\_UPA.  
DR Pfam; PF00051; Kringle; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM0020; Tryp\_Spc; 1.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS00070; KRINGLE\_2; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
DR PIRSF; PIRSF001144; Urk plasm act; 1.  
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.  
SQ SEQUENCE 433 AA; 48375 MW; 65E84F36415549B0 CRC64;  
  
Query Match 87.6%; Score 1321; DB 6; Length 433;  
Best Local Similarity 87.3%; Pred. No. 4.1e-123;  
Matches 241; Conservative 15; Mismatches 20; Indels 0; Gaps 0;  
  
QY 1 KPSPPPEELKFCQCKTLRPRFKIIGGEFTTIENQPFAAIYRRHRGGSVTVCGSLIS 60  
Db 158 KPALPPGKLEFCQCKALRPRFKIIGGEFTTIENQPFAAIYRRHRGGSVTVCGSLIS 217  
  
QY 61 PCWVISATHCFIDYPKKEDYIVLGRSRLNSNTQGMKFEVENLIHKDYSADTLAHND 120  
Db 218 PCWVSATHCFINHQKEDYIVLGRSRLNSMTFGMKFEVEQLIHGYSADTLAHND 277  
  
QY 121 IALLKIRSKGRCAQPSRTIOTICLPSMYNDPFGTSCETITGFKENSTDYLYPEQLKMT 180  
Db 278 IALLKILSNNGCAQPSRSITICLPWNADPNFGTSCETITGFKENSTDYLYPEQLKMT 337  
  
QY 181 VVKLISHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQDGGGLVCSLQGRMTLTGIVS 240  
Db 338 VVKLVSYQECQPHYGVSEVTTKMLCAADPOWKTDSCQDGGGLVCSVQGRMTLTGIVS 397  
  
RESULT 2  
Q8MHY7 PRELIMINARY; PRT; 433 AA.  
AC Q8MHY7;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Urokinase-type plasminogen activator.  
GN UROKINASE.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX Sugiki M., Yoshida E., Anai K., Maruyama M.;  
RA submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Iano W., Watanabe M.;  
RT "Oryctolagus cuniculus urokinase-type plasminogen activator, mRNA,  
RT complete cds."  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  
DR EMBL; AY029517; AAK40239.1; -.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0016301; F:kinase activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:000508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR009003; Cys Ser trypsin.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR001254; Peptidase S1.  
DR InterPro; IPR001314; Peptidase S1A.  
DR InterPro; IPR008293; Pept\_S1A\_UPA.  
DR Pfam; PF00051; Kringle; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM0020; Tryp\_Spc; 1.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS00070; KRINGLE\_2; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
DR PIRSF; PIRSF001144; Urk plasm act; 1.  
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.  
SQ SEQUENCE 433 AA; 48444 MW; 6DD35A371010A6EE CRC64;  
  
Query Match 87.3%; Score 1316; DB 6; Length 433;  
Best Local Similarity 87.0%; Pred. No. 1.3e-122;  
Matches 240; Conservative 15; Mismatches 21; Indels 0; Gaps 0;  
  
QY 1 KPSPPPEELKFCQCKTLRPRFKIIGGEFTTIENQPFAAIYRRHRGGSVTVCGSLIS 60  
Db 158 KPALPPGKLEFCQCKALRPRFKIIGGEFTTIENQPFAAIYRRHRGGSVTVCGSLIS 217  
  
QY 61 PCWVISATHCFIDYPKKEDYIVLGRSRLNSNTQGMKFEVENLIHKDYSADTLAHND 120  
Db 218 PCWVSATHCFINHQKEDYIVLGRSRLNSMTFGMKFEVEQLIHGYSADTLAHND 277  
  
QY 121 IALLKIRSKGRCAQPSRTIOTICLPSMYNDPFGTSCETITGFKENSTDYLYPEQLKMT 180  
Db 278 IALLKILSNNGCAQPSRSITICLPWNADPNFGTSCETITGFKENSTDYLYPEQLKMT 337  
  
QY 181 VVKLISHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQDGGGLVCSLQGRMTLTGIVS 240  
Db 338 VVKLVSYQECQPHYGVSEVTTKMLCAADPOWKTDSCQDGGGLVCSVQGRMTLTGIVS 397  
  
QY 241 WGRGCAKDKPGVTVRVSHFLPWIRSHTKENGLAL 276

```
Db 398 WGGCALKNKGVTYVRSFLFWIRSHIGEENGLAL 433
RESULT 3
Q9XT70
ID Q9XT70 PRELIMINARY; PRT; 214 AA.
AC Q9XT70;
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Urokinase-type plasminogen activator (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Vin J., Idell S.;
RT "Partial mRNA of rabbit uPA.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF097647; AAD39351.1; -.
DR HSSP; P00749; 1EJN.
DR MEROPS; S01.231; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008301; F:kinase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009001; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; kringle; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00089; trypsin; 1.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SW00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Serine protease.
FT NON_TER 214 214
SQ SEQUENCE 214 AA; 24314 MW; 69975C41C32B0D7E CRC64;

Query Match 44.4%; Score 670; DB 6; Length 214;
Best Local Similarity 84.2%; Pred. NO. 1.5e-58;
Matches 123; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 1 KPSSPEELKFCQGKTLSPRKILGGFTTIENQWPAAYRHRGGSVTVVCGSLIS 60
DB 69 KPALPPGKLEFCQGKALPRKILGGFTTIENQWPAAYRHRGGSVTVVCGSLIS 128
QY 61 PCWVSAATHCFIDYPKEDYIVYLGSRSLNSNTQGMKEPEVENLILHKDYSADTLAHND 120
DB 129 PCWVSAATHCFIDYPKEDYIVYLGSRSLNSNTQGMKEPEVENLILHKDYSADTLAHND 188
QY 121 IALLKRSKEGRCQAQPSRIQICLP 146
DB 189 IALLKLSNNGCQAQPSRSIQICLP 214

RESULT 4
Q97587
ID Q97587 PRELIMINARY; PRT; 128 AA.
AC Q97587;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
```

```
DE Urokinase (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand White;
RX MEDLINE=99057575; PubMed=9837780;
RA Reno C., Boykiw R., Martinez M.L., Hart D.A.;
RT "Temporal alterations in mRNA levels for proteinases and inhibitors
and their potential regulators in the healing medial collateral
ligament.";
RL Blochem. Biophys. Res. Commun. 252:757-763(1998).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AF069711; AAC95003.1; -.
DR HSSP; P00749; 1EJN.
DR MEROPS; S01.231; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
KW Hydrolase; Kinase; Protease; Serine protease.
FT NON_TER 1 1
FT NON_TER 128 128
SQ SEQUENCE 128 AA; 14328 MW; 1BC7ED30E071A06D CRC64;

Query Match 38.8%; Score 585; DB 6; Length 128;
Best Local Similarity 84.4%; Pred. NO. 2.3e-50;
Matches 108; Conservative 6; Mismatches 14; Indels 0; Gaps 0

QY 37 WFAAIYRHRGGSVTVVCGSLISPCWVSAATHCFIDYPKEDYIVYLGSRSLNSNTQGE 96
DB 1 WFAAIYRHRGGSVTVVCGSLISPCWVSAATHCFIDYPKEDYIVYLGSRSLNSNTQGE 60
QY 97 MKPEVENLILHKDYSADTLAHNDIALKLSRSGRCQAQPSRIQICLPSTMTNDPQFGT 156
DB 61 MKPEVEQLILHEGYRADTLAHNDIALKLSNNGCQAQPSRSIQICLPWKNADPNFGT 120
QY 157 SCEITGFG 164
DB 121 SCEITGFG 128

RESULT 5
Q9BU99
ID Q9BU99 PRELIMINARY; PRT; 516 AA.
AC Q9BU99;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to plasminogen activator, tissue.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; BC002795; AAH02795.1; -.

```

```

DR HSP; P00750; IASH.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF_1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS02440; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 516 AA; 57370 MW; BAB31901FDC96800 CRC64;

Query Match 38.7%; Score 583; DB 4; Length 516;
Best Local Similarity 44.9%; Pred. No. 2.2e-49;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-QKTLRPFKILGGEFTTIENQWFAAIYRHH-RGGSVTVVCGSLISPCWVISATHC 70
DB 253 CGLRQSQQFRKIGGLFADIAHPQAAIFAKHRRSPGERFLCGGILISSCWILSAHC 312

QY 71 FIDYPKEDYIVYLGSRSLNNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSK 130
DB 313 FQERFPFPHLTVLIGRTYRVVPGEEBQKEVEKYIVHKEFDDDT--YNDIALLOLKSDS 370

QY 131 GRCAQSRITQICLSMYNDQFGTSCITGFGKENSIDYLPQOLKMTVVKLISHREC 190
DB 371 SRCAQSSVVRTVCLPPADQLPDWTECELSGKGHEALSPFYSERLKEAHVLYPSRC 430

QY 191 QQPHYGSEVTTKMLCAAD-----PQWKT-DSQQSDSGGPLVCSLQGRMTLTIGVSWGRG 244
DB 431 TSQHLLNRTVTDNMLCAGDTRSGGPQANLHDAQQDGGPLVCLNDGRMTLVGIISWGLG 490

QY 245 CALKDKPGVYTRVSHLPWIRSHTK 269
DB 491 CGQKDPGVYTKVTNYLDWIRDNR 515

RESULT 6
Q86YK8 PRELIMINARY; PRT; 562 AA.
AC Q86YK8
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Tissue plasminogen activator.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Liu Y., Xu L., Zeng Y., He X.;
RT "cDNA of tissue plasminogen activator";
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY221101; AAC34406.1; -.
DR GO; GO:000576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibronectin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS02440; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
SQ SEQUENCE 562 AA; 62902 MW; 837D98392F65DD1F CRC64;

Query Match 38.7%; Score 583; DB 4; Length 562;
Best Local Similarity 44.9%; Pred. No. 2.4e-49;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-QKTLRPFKILGGEFTTIENQWFAAIYRHH-RGGSVTVVCGSLISPCWVISATHC 70
DB 299 CGLRQSQQFRKIGGLFADIAHPQAAIFAKHRRSPGERFLCGGILISSCWILSAHC 358

QY 71 FIDYPKEDYIVYLGSRSLNNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSK 130
DB 359 FQERFPFPHLTVLIGRTYRVVPGEEBQKEVEKYIVHKEFDDDT--YNDIALLOLKSDS 416

QY 131 GRCAQSRITQICLSMYNDQFGTSCITGFGKENSIDYLPQOLKMTVVKLISHREC 190
DB 417 SRCAQSSVVRTVCLPPADQLPDWTECELSGKGHEALSPFYSERLKEAHVLYPSRC 476

QY 191 QQPHYGSEVTTKMLCAAD-----PQWKT-DSQQSDSGGPLVCSLQGRMTLTIGVSWGRG 244
DB 477 TSQHLLNRTVTDNMLCAGDTRSGGPQANLHDAQQDGGPLVCLNDGRMTLVGIISWGLG 536

QY 245 CALKDKPGVYTRVSHLPWIRSHTK 269
DB 537 CGQKDPGVYTKVTNYLDWIRDNR 561

RESULT 7
Q9BZW1 PRELIMINARY; PRT; 395 AA.
AC Q9BZW1
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Neonatal thrombolytic agent alpha-form (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Dou D.;
RT "A brain-type plasminogen activator";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

```

```
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF260845; AAK11956.1; -.
DR HSSP; P00750; 1PK2.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_ser_trypsin.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
FT NON TER 395
SQ SEQUENCE 395 AA; 44323 MW; 3FBD4A2F0B7C11C8 CRC64;

Query Match 38.0%; Score 573; DB 4; Length 395;
Best Local Similarity 44.2%; Pred. No. 1.5e-48;
Matches 117; Conservative 39; Mismatches 99; Indels 10; Gaps 5;

QY 13 CG-QXTLRPRKLIIGETTTIENQFWFAIYRRH-RGGSVTVVCGSLISPCWISATHC 70
DB 132 CGLRQYKQPKRIKGLYADITSHPWQAAIFVKNRSPGCRFLCGGILISSCWLSAAHC 191

QY 71 FIDYPKKEDYIVVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHHNDIALLKIRSK 130
DB 192 FQERPPPHLTVILGRTVVRVPGEEQKFECEKIVHKEFDDT--YNDIALQLKSDS 249

QY 131 GRCAQPSRTIQTICLPSMYNDPQFTSCETIGFGKENSTDYLYPEQLKMTVVKLISHREC 190
DB 250 SRCAQESSVVRTVCLPFPADLQLPDWTECELSGKGHEALSPFYSERLKEAHRVLYPSSRC 309

QY 191 QQPHYGSEVTTKMLCAADP-----PWKT-DSQCQSGGGLVCSLQGSMTLTGIVSWG 244
DB 310 TSQHLLNRTVTDNMLCAGDTRSGGQANLHDACQSGGGLVCLNDGNTLVGLISWGLG 369

QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
DB 370 CGQKDVGVYTKVNYLDWIRDNR 394

RESULT 8
Q8SQ23 ID Q8SQ23 PRELIMINARY; PRT; 562 AA.
AC Q8SQ23;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE T-plasminogen activator.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Enamel organ;

RA Ding Y., Xue J., Bartlett J.D.;
RT "T-plasminogen activator in tooth tissues.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AF364605; AAM00297.1; -.
DR HSSP; P00761; 1ANI.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_ser_trypsin.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRODOM; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 562 AA; 63668 MW; F9B6B4C77CB101E8 CRC64;

Query Match 36.7%; Score 553; DB 6; Length 562;
Best Local Similarity 41.9%; Pred. No. 2.4e-46;
Matches 111; Conservative 43; Mismatches 101; Indels 10; Gaps

QY 13 CGQKTLR-PRKLIIGETTTIENQFWFAIY-RRHGGSVTVVCGSLISPCWISATHC 70
DB 299 CGLRQYKQPKRIKGLYADITSHPWQAAIFVKNRSPGCRFLCGGILISSCWLSAAHC 35

QY 71 FIDYPKKEDYIVVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHHNDIALLKIRSK 13
DB 359 FQERPPPHVVRVVLGRTVLRVPGEEQAEVEKYIVHKEFDDT--YNDIALQLKSDS 41

QY 131 GRCAQPSRTIQTICLPSMYNDPQFTSCETIGFGKENSTDYLYPEQLKMTVVKLISHREC 19
DB 417 LTCAQESDAVRTVCLPFPANLQLPDWTECELSGKGHEALSPFYSERLKEAHRVLYPSSRC 47

QY 191 QQPHYGSEVTTKMLCAADP-----KTDSCQSGGGLVCSLQGSMTLTGIVSWG 24-
DB 477 TSXHLFNKTIITNNMLCAGDTRSGGDVANLHDACQSGGGLVCMKGNHTLVGLISWGLG 53

QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
DB 537 CGQKDVGVYTKVNYLDWIRDNR 561

RESULT 9
Q8MKB1 ID Q8MKB1 PRELIMINARY; PRT; 564 AA.
AC Q8MKB1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
```



DE Tissue-type plasminogen activator.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Sugiki M., Yoshida E., Anai K., Matuyama M.;  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.  
DR EMBL; AY029518; AAK40240.1; -.  
DR GO; GO:0005576; C:extracellular; IEA.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR000883; Fibrinctn1.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00008; EGF; 1.  
DR Pfam; PF00039; fn1; 1.  
DR Pfam; PF00051; kringle; 2.  
DR PRINTS; PR00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR PRODOM; PD000395; Kringle; 2.  
DR SMART; SM00181; EGF; 1.  
DR SMART; SM00058; FNL; 1.  
DR SMART; SM00130; KR; 2.  
DR SMART; SM00020; TRYPSIN; 1.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS01186; EGF\_2; 1.  
DR PROSITE; PS01253; FIBRONECTIN\_1; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 2.  
DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
DR EGF-like domain; Glycoprotein; Hydrolase; Protease;  
KW Serine protease.  
SQ SEQUENCE 564 AA; 62726 MW; 459DBBAC6D4A937C CRC64;  
  
Query Match 36.2%; Score 545.5; DB 6; Length 564;  
Best Local Similarity 43.1%; Pred. No. 1.3e-45;  
Matches 115; Conservative 36; Mismatches 103; Indels 13; Gaps 4;  
  
Qy 12 QCGQKTLR----PRKTIIGBFTTIENQWPAATV-RRHGGSVTVYVCGSLISPCWVTS 66  
Db 297 QCATGLRQDKQKQPRKIGSLFDITAHFWAALFTNNRSPGRLFCGILLNCWVLS 356  
  
Qy 67 ATHCFIDYPKRSDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALTKI 126  
Db 357 AARCFLEPFQOKLAVILGRVYPLVSAEEQIFVEQFILHERFDEGT--YNDIALKL 414  
  
Qy 127 RSKEGRCAQPSRTIOTICLPMSYNDPQGTSCETIGFKXENSDIYLPQKMTVVVKLIS 186  
Db 415 KSTSGSCAQESQAVRLVCLPDASLQLPDWTSCSLGSGYKHEEFPVPSQLKEAHLVLP 474  
  
Qy 187 HRECCQPHYGVSEVTMLCAADPW-----KTDSCQSGSGPLVCSQGRMLTGLVS 240  
Db 475 SSRCTPQQLKNTVTNMLCAGDTRSGAQVNLHDAQCGSGGGLVQMTDGHMTLIGITS 534  
  
Qy 241 WGRGCALKDKPGVYTVSHFLPWIRSH 267  
Db 535 WGLGCGQKQDVGVTYTKVNYLQIQH 561

RESULT 10

Q8VCS4  
ID Q8VCS4  
AC Q8VCS4;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strusberg R.;  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  
DR EMBL; BC019376; AAHL9376.1; -.  
DR HSP; P00761; IAN1.  
DR GO; GO:0005576; C:extracellular; IEA.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR00742; EGF\_2.  
DR InterPro; IPR006209; EGF-like.  
DR InterPro; IPR000883; Fibrinctn1.  
DR InterPro; IPR000562; FN\_Type\_II.  
DR InterPro; IPR004210; IEGF.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00008; EGF; 2.  
DR Pfam; PF00039; fn1; 1.  
DR Pfam; PF00040; fn2; 1.  
DR Pfam; PF00051; kringle; 1.  
DR PRINTS; PR00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00013; FNTYPEII.  
DR PRINTS; PR00018; KRINGLE.  
DR PRODOM; PD000395; FN\_Type\_II; 1.  
DR PRODOM; PD000395; Kringle; 1.  
DR SMART; SM00181; EGF; 2.  
DR SMART; SM00059; FN2; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00020; TRYPSIN; 1.  
DR PROSITE; PS00022; EGF\_1; 2.  
DR PROSITE; PS01186; EGF\_2; 1.  
DR PROSITE; PS01253; FIBRONECTIN\_1; 1.  
DR PROSITE; PS00023; FIBRONECTIN\_2; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS00070; KRINGLE\_2; 1.  
DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
DR Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;  
KW Kringle; Protease; Serine protease.  
SQ SEQUENCE 653 AA; 70553 MW; FE18D90174ED6FDD CRC64;  
  
Query Match 33.0%; Score 497.5; DB 11; Length 653;  
Best Local Similarity 38.5%; Pred. No. 9.9e-41;  
Matches 105; Conservative 44; Mismatches 109; Indels 15; Gaps 5  
  
Qy 2 PSSPPEELKFCQGGK-----TLRPRFKIIGBFTTIENQWPAATVRRHGGSVTVYVCGS 5  
Db 382 PESAP-AVRPTCGKHKRKTFLRPR--IIGSSSLPGSHFWLAAY---IGNS---FCAG 4  
  
Qy 57 SLISPCWVISATHCFIDYPKRSDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLA 1  
Db 433 SLVHTCWVVSAAHCFANSPRDSITVVLGQHFFNRITDVTQTFGIEKYYVYTLVSFNP 4

QY 117 HNDIALKIRSKGRCACQPSRTTICLPSMYNDPQGTSCITGFGKNSDYLYPEQ 176  
DB 493 NH-DLVLRLLKKGGRCACVRFQVQICLPAGSGSPTGHCQIAGHMDENVSSYNS 551  
QY 177 LKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGGGLVCSLQGRWTLT 236  
DB 552 LLEALVPLVADKKCSPEVYGADISPNMLCAGYFDCKSDACQDGGGGLVCKNGVAYLY 611  
QY 237 GIVSGRCALKDKPGVTVTRVSHLPWIRSHK 269  
DB 612 GIISWGCGRLNKGVTIRVANVYDWINDR 644

RESULT 11

Q95M89 PRELIMINARY; PRT; 103 AA.  
AC Q95M89  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Plasmidogen activator urokinase (Fragment).  
GN PLAU.  
OS Equus caballus (Horse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
OX NCBI\_TaxID=9796;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21314992; PubMed=11421942;  
RA Shubitski D.M., Vanta P.J., Douglass C.L., Zhou R.-X., Ewart S.L.;  
RT "Polymorphism identification within 50 equine gene-specific sequence tagged sites."; 78-78 (2001).  
RL Annu. Genet. 32:78-78 (2001).  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
DR EMBL; AY008806; AAK14840.1; JOINED.  
DR EMBL; AY008803; AAK14840.1;  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0016301; E:kinase activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR009003; Cys Ser trypsin.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS0240; TRYPsin DOM; 1.  
DR PROSITE; PS00134; TRYPsin HIS; 1.  
KW Hydrolase; Kinase; Protease; Serine protease.  
FT NON\_TER 1  
FT NON\_TER 103  
SQ SEQUENCE 103 AA; 11525 MW; 0B739514F6331180 CRC64;

Query Match 32.8%; Score 495; DB 6; Length 103;  
Best Local Similarity 87.4%; Pred. No. 1.6e-41;  
Matches 90; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 37 WFAAIYRRHGGSVTVVCGSLISPCWVTSATHCFIDYPKKEDYIVYLGSRRLNSNTGGE 96  
DB 1 WFAAIYRRHGGSVTVVCGSLISPCWVLSATHCFINTPKKEDYIVYLGSRRLNSTSGE 60  
QY 97 MKFEVENLILHKDYSADTLAHENDIALKIRSKGRCACQPSRT 139  
DB 61 MKFEVEKLILHEDYSADTLAHENDIALKIRSKSTGCAQPSRS 103

RESULT 12

Q96EF3 PRELIMINARY; PRT; 300 AA.  
AC Q96EF3  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein (Coagulation factor XII) (Hageman factor).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RA Strausberg R.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBSJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Kalline N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,  
RA Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,  
RA Phelan M., Farmer A.;  
RT "Cloning of human full-length cDNAs in BD Creator(TM) System Donor vector.";  
RL Submitted (MAY-2003) to the EMBL/GenBank/DBSJ databases.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
DR EMBL; BC012390; AAH12390.1; -;  
DR EMBL; BT007350; AAP36014.1; -;  
DR HSP; P00761; IAN1.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR009003; Cys Ser trypsin.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS0240; TRYPsin DOM; 1.  
DR PROSITE; PS00134; TRYPsin HIS; 1.  
DR PROSITE; PS00135; TRYPsin SER; 1.  
KW Hypothetical protein; Hydrolase; Protease; Serine protease.  
SQ SEQUENCE 300 AA; 32216 MW; FFC2BDF9382F636A CRC64;

Query Match 32.8%; Score 495; DB 4; Length 300;  
Best Local Similarity 38.5%; Pred. No. 6.5e-41;  
Matches 107; Conservative 44; Mismatches 109; Indels 18; Gaps

QY 1 KPSSPPEELK---FQCGQ---KTLRPRFKIIGGEFTTIENQWFAIYRRHGGSVTVVC 54  
DB 29 KREQPSLTRNGPLSCGQLRKLSLSMTRVVGLVALRGHPIYALYNGHS-----FC 82  
QY 55 GGSLLISPCWVTSATHCFIDYPKKEDYIVYLGSRRLNSNTGEMKFEVENLILHKDYSADT 11  
DB 83 AGSLIAPCWVLTAAHCLQDRPAPEDLTIVYLGSRRLNSNTGEMKFEVENLILHKDYSADT 14  
QY 115 LAHNDIALKIR-SKEGRCAQPSRTIQTICLPSMYNDPQGTSCITGFGKNSDYL 17  
DB 141 VSTQHDALLRLQEDADGSCALLSPYQVCLPSSGAARFSETTLCOVAGHGQFEGAEY 20  
QY 174 PEQLKMTVVKLIISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGGGLVCSLQ-- 23  
DB 201 ASFLQEAQVPLSLERCSAPDVHGSSILPGLMCAFLGEGTDACQDGGGGLVCSLQ-- 26  
QY 232 -RWTLGIYVSGRCALKDKPGVTVTRVSHLPWIRSHK 268  
DB 261 RLTLQGIISWGCGRLNKGVTIRVANVYDWINDR 298

RESULT 13

Q800Y7 PRELIMINARY; PRT; 540 AA.  
AC Q800Y7  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hepatocyte growth factor activator (Fragment).  
OS Meleagris gallopavo (Common turkey).



```
QY 115 LAHNDJALLKIR-SKEGRCAQPSRTIOTICLPMSYNDPQGTSCETITGFGKNSDYL 173
DB 456 VSYQHDALLR-LEDADGSCALLSPYQVPCVLPAGARPSETTLCQVAGCGHFEAGAEY 515
QY 174 PEOLKQVTVKLISHRECQPHYGVSEVTKMLCAADPQWKTDSCQDGGGGLVCSLQ-- 231
DB 516 ASFLQEAQVPLSLRCSAPDVHGSSILPGMLCAGFLEGGTDACQDGGGGLVCEQAAE 575
QY 232 -RMTLTGTVSWGRGCAKDKPGVTVRVSHFLPWIRSH 268
DB 576 RLTLQGLISWGGCGDRNKPQVTVDAVYLAWEHT 613

RESULT 15
O97507 PRELIMINARY; PRT; 616 AA.
AC O97507;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE FXII.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]_TaxID=9823;
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Takahashi T., Kihara T.;
RT "Porcine liver factor XII.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AB022426; BAA37148.1; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.211; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibnctn1.
DR InterPro; IPR000562; FN Type_II.
DR InterPro; IPR062210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN Type II; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00032; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01293; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
```

```
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 616 AA; 58012 MW; 4CSF3D71EBBDIA9 CRC64;

Query Match 31.0%; Score 467; DB 6; Length 616;
Best Local Similarity 37.8%; Pred. No. 1e-37;
Matches 107; Conservative 46; Mismatches 102; Indels 28; Gaps 9

QY 4 SPPEL-----KFOCGOKTLRPRF---KIIGSEFTTIENQWFAALYRHRGGSVTV 52
DB 342 APPEQRGLPSAGLVGGQR-LKRLSSLNRIUGLVALPGAHPYIALYWGQN----- 394
QY 53 VCGSLISPCWVISATHCFIDYPKEDYIVYGRSLNNTQGMKFVENILHXDYSA 112
DB 395 FCAGSLIAPCWLTAAHCLQNPAPBELTIVLQDRHNSCEQCQTAVRSYRLHESYSP 454
QY 113 DTLAHNDJALLKIR-SKEGRCAQPSRTIOTICLP--SMYNDPQGTSCETITGFGKNS 168
DB 455 KTYQH--DLALVRLKETADGCCAHPSPFPQVCLPRSVASSAEPE-GALCEVAGWGHQFE 511
QY 169 TDLYLPEQLKQVTVKLISHRECQPHYGVSEVTKMLCAADPQWKTDSCQDGGGGLVCSLQ-- 227
DB 512 GAEYSYSLQEAQVPLSLRCSADVHGAAFTPGMLCAGFLEGGTDACQDGGGGLVCE 571
QY 228 --SLOGRMTLTGTVSWGRGCAKDKPGVTVRVSHFLPWIRSH 268
DB 572 DETAERQLVLRGIVSWGGCGDRNKPQVTVDAVYLAWEHT 614

RESULT 16
O9Y1V3 PRELIMINARY; PRT; 866 AA.
AC O9Y1V3;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Tunicate retinoic acid-inducible modular protease precursor.
GN TRAMP.
OS Polyandrocarpa misakiensis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Styelidae; Polyandrocarpa.
OX NCBI_TaxID=7723;
RN [1]_TaxID=7723;
RP SEQUENCE FROM N.A.
RC STRAIN=white spot;
RX MEDLINE=39423646; PubMed=10491255;
RA Ohashi M., Kawamura K., Fujii N., Yubisui T., Fujiwara S.;
RT "A retinoic acid-inducible modular protease in budding ascidians.";
RL Dev. Biol. 214:38-45(1999).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AB030007; BAA82522.1; -.
DR HSSP; P00763; IDPO.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0005044; F:scavenger receptor activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00057; ldl_recept_a; 3.
DR Pfam; PF00024; PAN_1.
DR Pfam; PF00530; SECR; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00261; LDLRECEPTOR.
```

```

DR PRINTS: PR00259; SPERACTRCPT.
DR SMART; SM00192; LDLA_3.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00202; SR; 2.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS01209; LDLA_1; 3.
DR PROSITE; PS00688; LDLA_2; 3.
DR PROSITE; PS0287; SRCR_2; 2.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Protease; Serine protease; Signal.
FT SIGNAL
SQ SEQUENCE 868 AA; 97660 MW; F71462865F36A6CA CRC64;

Query Match
Best Local Similarity 30.9%; Score 463.5; DB 5; Length 868;
Matches 107; Conservative 44; Mismatches 103; Indels 21; Gaps 10;

DR 2 PSSPPEELKFOGQKTL-----RPRFKIIGGEFTTIENQFWFAAIYRRHGGSVTVVCGGS 57
DB 601 PTPPPMP-----ECGRKPVIEAPLPTARIIVGSGSGTEPHEWFOAGIWL-----PWTVMCGGS 652
QY 58 LISPQWISATCFI-DYPKEDYIVYLGSRSLNNTQGMKFEVENILHKDYSADTLA 116
DB 653 LHPCWLTAAHCFVREYFIR-DYIRLGDHITGVDDTEFLKAEIKH-DYIVTT-- 708
QY 117 HNDIALKIRKSGRCAQPSRTIQTICLPSMYNDPQFGTSCBITGFGKENSTDY-LYPE 175
DB 709 KENDIALRIENDARECATITPEVQTVCLPKSSQFDAKTICEVTGWGKGATAVRYPV 768
QY 176 QLKXMTVVKLISHRECOOPHYGSEVTKMLCAADPQWKTDSCQDGGGGLVCSLOG--RM 233
DB 769 VLQAEIPLIANKCLRDSEY-TQLGPTMFCAGYVTGKDKSCQDGGGGLVCSRDQSDRY 827
QY 234 TLTGIVSWGRCALKDKPGVYTRVSHFLFWIRSH 268
DB 828 YVWGVSWGNCAPKAPGAYKAVFIDWIEQMT 862

RESULT 17
ID Q35727 PRELIMINARY; PRT; 597 AA.
AC Q35727;
DT 01-JAN-1998 (T-EMBLrel. 05, Created)
DT 01-JUN-1998 (T-EMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Factor XII.
GN F12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Schloesser M., Schwager S., Engel W.;
RA Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; X9571; CA67891.1; -.
DR HSSP; P00760; IAQ7.
DR MEROPS; S01.211; -.
DR MGPI; MG1:1891012; F12.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; ENTPEPII.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000995; FN_Type_II; 1.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; K3; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS01253; FIBRONECTIN_2; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 597 AA; 65638 MW; F3AC07C37D0C0PBA CRC64;

Query Match
Best Local Similarity 29.9%; Score 451.5; DB 11; Length 597;
Matches 97; Conservative 47; Mismatches 103; Indels 15; Gaps 5

QY 13 CGQ---KTLRPRFKIIGGEFTTIENQFWFAAIYRRHGGSVTVVCGSLISPCWVISATH 69
DB 341 CGQFRKGLSSFMVVGVLVPGSHFYAALYGNV-----FCAGSLIAPCWLTAAH 394
QY 70 CFIDYPKEDYIVYLGSRSLNNTQGMKFEVENILHKDYSADTLAHENDIALKIR-S 128
DB 395 CLQNRPAPELTVVLGDRHNCSCWCQTLAVRSYELHEGFSITVQH--DLALLQLQS 452
QY 129 KEGRCAQPSRTIQTICLPSMYNDPQFGTSCBITGFGKENSTDYLYPEQLXMTVVKLISHR 188
DB 453 KTNSCALISPHVQPVCLPSGAAPPSETVLCEVAGWGHLGAEYSFTLQEAQVPFIALD 512
QY 189 ECQPHYGYSEVTKMLCAADPQWKTDSCQDGGGGLVCSLOGMTLTGIVSWGRC 245
DB 513 RCSNSNVHGDAILFGMLCAGFLGGTDACQDGGGLVCEGTAHQLTLRGVISWGSGC 572
QY 246 ALKDKPGVYTRVSHFLFWIRSH 267
DB 573 GDRNKGVTVDVANYLAWIQKH 594

RESULT 18
Q80YCS PRELIMINARY; PRT; 609 AA.
ID Q80YCS;
AC Q80YCS;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Similar to coagulation factor XII (Hageman factor) (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Liver;
RA Strausberg R.;

```



```

BL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AF536382; AAN04055.1; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; tryp_Spc; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 284 AA; 30110 MW; FDF3F1750D569978 CRC64;

Query Match 28.9%; Score 435.5; DB 4; Length 284;
Best Local Similarity 38.4%; Pred. No. 5.2e-35;
Matches 103; Conservative 35; Mismatches 97; Indels 33; Gaps 8;

QY 13 CQOKTLRPFKIGGEFTTIENQPFALYRHRGGSVTVVCGSLSPCWVISAHCFFI 72
DB 28 CQPRMSSR--IVGGRDGRGGEWPCASI--QHPG---ARVCGGSLIAPQWVLTAAHCFF 80

QY 73 DYPKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHHNDIALLKIRSKGR 132
DB 81 RRALPARYRVLGALRLGTSPTLSVPVRVLLPPDYSED--GARGDLALLQLR---R 134

QY 133 CQAPSRTIQTICLPMSYNDPQGTSCITGFGKENSITLYP-----EQKMTVVKLI 185
DB 135 PVPLSARVQVCLPVEGARPPPTPCRVTVGWS-----LRPGVLPWRPLQGVVRVPL 188

QY 186 SHRECOQPHYGVSEVTTK-----MLCAADPQWKTDCSQGDSGGLVCSLQGRMTLTGI 238
DB 189 DSRTDGLYHVGADVPQASRIVLPGLSCAGYQGHKDAQGDSGGPLTCLQSGSNVLGV 248

QY 239 VSWGRGALKDQGVTVTRVSHFLPWIRS 266
DB 249 VSWGKCALPNRPGVYTSVATYSPWIOA 276

RESULT 21
Q9BK47 PRELIMINARY; PRT; 267 AA.
ID Q9BK47
AC Q9BK47;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sea star regeneration-associated protease SRAP.
OS Luidia foliolata.
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;
OC Asterozoa; Valvatacea; Paxillozoa; Luidiida; Luidia.
OX NCBI_TaxID=105861;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21100442; PubMed=11179669;
RA Vickers M.C.L., Vickers M.S., McClintock J.B., Amsler C.D.;
RT "Utilization of a novel deuterostome model for the study of
RT regeneration Genetics: Molecular cloning of genes that are
RT differentially expressed during early stages of larval sea star
RT regeneration."
RL Gene 262:73-80(2001).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AF312826; AAK15274.1; -.
DR HSP; P00763; IDPO.
DR Pfam; PF00089; trypsin; 1.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.

Q966V4 PRELIMINARY; PRT; 505 AA.
ID Q966V4
AC Q966V4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Proacrosin.
OS ACR.
OC Halocynthia roretzi (Sea squirt).
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Pyridae; Halocynthia.
OX NCBI_TaxID=7729;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21326076; PubMed=1113341;
RA Kodama E., Baba T., Yokosawa H., Sawada H.;
RT "cDNA Cloning and Functional Analysis of Ascidian Sperm Proacrosin."
RL J. Biol. Chem. 276:24594-24600(2001).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.
DR EMBL; AB052635; BAB60718.1; -.
DR HSP; P00761; IAN1.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS01180; CUB; 2.
```



```
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 505 AA; 55002 MW; 79A1A917CE1D9334 CRC64;

Query Match      28.2%; Score 425.5; DB 5; Length 505;
Best Local Similarity 35.5%; Pred. No. 1.1e-33;
Matches 99; Conservative 47; Mismatches 92; Indels 41; Gaps 11;

QY 13 CGQKTLPRF-----KIIGGETTINQWFAAIYRRHGGSVTVVCGSLISPCWVI 65
DB 21 CG---LPRLOSALTIGVIGEMAKLGEFPWQAFLYKH-----VQVCGTIDITWIL 72
QY 66 SATKCF-----IDYKKEDVIVYLG-RSRNLNSTQGEKMFVENLILHKDYADTLAAH 118
DB 73 SAAGCFDPMYNLOSIIKKEDALIRVADLDDTDDDEGMTFEVXDIIIEHQYNRQTF--D 130
QY 119 NDIALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFQTSCEITGFGKENSTDYL---YPE 175
DB 131 NDIMIEILGS-----ITVGPVQACIPGANDAVADGTKLISWG--DTQDHYHNRWPD 184
QY 176 QLKMTVVKLISHRECCQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVC-----SLQ 230
DB 185 KLQKAQVEVFARAQCLATY---PESTENNICAGLRTGGIDSCQDGGPLACPFPTNTAQ 241
QY 231 GRVTLTGIVSGRGCAKDKQGVTVRSHFLPWIRSHTK 269
DB 242 PTFPLQGVSGRGCAKDKGFGGVTEVRKYSWIANITQ 280

RESULT 23
Q8K0D2 PRELIMINARY; PRT; 517 AA.
AC Q8K0D2;
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TremBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RC TISSUE=Kidney;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC031775; AAH31775.1; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0005508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR0018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.

DR PROSITE; PS00022; EGF_1; 3.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
KW Kringle; Protease; Serine protease.
SQ SEQUENCE 517 AA; 57326 MW; 3855A42035A5EA59 CRC64;

Query Match      28.2%; Score 425; DB 11; Length 517;
Best Local Similarity 38.7%; Pred. No. 1.3e-33;
Matches 101; Conservative 36; Mismatches 102; Indels 22; Gaps 9.

QY 23 KIIGGETTINQWFAAIY-----RHRGGSVTVVCGSLISPCWVISATHCFIDYP 75
DB 270 RYGGFKSTAGKHPQVSLQTSPLTTSMPQG---HFCGGALHPCWVLTAAHC-TDIN 324
QY 76 KKEEDVYVIGRSLNSNTQGEKMFVENLILHKDYADTLAHHNDIALLKIRSEGRCAQ 135
DB 325 THLKV-V-LGDQDLKTESHEQTFVEKILKYSQYNERDEIPHNDIALLKLPVGHCA 383
QY 136 PERTIQTICLPSMYNDP-QFGTSCEITGFGKENSTDYLPEQLKMTVVKLISHRECCQPH 194
DB 384 ESRVYKTVCLPS---DPFPGSTECHISGWVETGSE--GSRQLLDAKVKLIANPLNSRQ 438
QY 195 YGVSEVTTKMLCAADPQWK-TDSCQDGGPLVCSLQGRMTLTGIVSGRGCAKDKQGV 253
DB 439 LYDHTTDDSMICAGNLQKPGSDTCQDGGPLTCCKDGTYYVYGVISWQEGCG--KKPGV 496
QY 254 YTRVSHFLPWIRSHTKENGL 274
DB 497 YIQVTKFLNWKTTMREAGL 517

RESULT 24
Q99JC8 PRELIMINARY; PRT; 761 AA.
AC Q99JC8;
DT 01-JUN-2001 (TremBLrel. 17, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Neurotrypsin.
GN NT.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Hintsch G., Sonderegger P.;
RL "Cloning and sequencing of the cDNA encoding rat neurotrypsin.";
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ311671; CAC35028.2; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0005044; F:scavenger receptor activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00530; SRCR; 3.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00258; SPERACTROPT.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00202; SR; 3.
```





Db 77 --PRVWPESEVILGALSLOVRSGHELLVPLVRLVLLPPDYSEDEA--RGLDALLQLRHP 132  
Qy 130 EGRCAPQERTTQICLPSMYNDPQGTSCETITGKGNSTDYLP-----EOLKMTVV 182  
Db 133 ---VSLSTRIQVCLPAPGHPGPPGSPCWITGWS-----LSPGVLPKRPLQGVV 182  
Qy 183 KLISHRECOQPHYVSE-----VTKMLCAADPQWKTSCQDGGPLVCSLQGRMTL 235  
Db 183 PLLSRACDRLYHVGANVPOGERIVLPONLCAGYRRGHKDACQDGGPLTCMESGEHWL 242  
Qy 236 TGIVSWGRGCGALKDKPGVYTRVSHFLPWIR 266  
Db 243 VGVVSWGKGCALPNRPGYTNVAKYSPWIA 273

RESULT 27  
Q9BYE1 PRELIMINARY; PRT; 537 AA.  
AC Q9BYE1; 01-JUN-2001 (TRENBLrel. 17, Created)  
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE Mosaic serine protease.  
GN MSPS.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RX MEDLINE=21167393; PubMed=11267681;  
RA Kim D.R., Sharmin S., Inoue M., Kido H.;  
RT "Cloning and expression of novel mosaic serine proteases with and  
RT without a transmembrane domain from human lung."  
RL Biochim. Biophys. Acta 1518:204-209 (2001).  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
DR EMBL; AB048797; BAB39742.1; -;  
DR HSSP; P00763; LDPO.  
DR MEROPS; S01.087; -;  
DR GO; GO:0016021; C:integral to membrane; NAS.  
DR GO; GO:0008508; P:proteolysis and peptidolysis; NAS.  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR002172; LDL\_receptor\_A.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR InterPro; IPR001190; Srcr\_receptor.  
DR Pfam; PF00089; ldl\_recept\_a; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00192; LDL1; 1.  
DR SMART; SM00202; SR; 1.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS02087; SRCR 2; 1.  
DR PROSITE; PS02040; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolyase; Protease; Serine protease.  
SQ SEQUENCE 537 AA; 58102 MW; A39FF4E8816DAECF CRC64;

Query Match 27.7%; Score 418; DB 4; Length 537;  
Best Local Similarity 36.7%; Pred. No. 6.6e-33;  
Matches 97; Conservative 38; Mismatches 105; Indels 24; Gaps 8;  
Qy 13 CGQKTLRPFKIIGGFTTIENQPPAAIYRRHGGSVTVVCGGSLISPCWVISAHCIFI 72  
Db 287 CGLRAMTGR--IVGGALASDKWPQVSL---HFG--TTHICGGTLIDAQWVLTAAHCFF 339  
Qy 73 DYPKK--EDYIVYLGSRSLNSNTQGMKFEVENLIHKDYSADTLAHHNDIALKIRSK 130  
Db 340 VTREKVLGKWKVAGTNSLHQLPEAA---SIAIINSNYTDE--EDDYDIALM----- 388  
Qy 131 GRCAQP---SRTIQTCLPSMYNDPQGTSCETITGKGNSTDYLPQKMTVVKLISH 187

Db 389 -RLSKPLTSLSAHIHPACLPFMHGGTFLSLNETCMTWITFGKTRTDDTKTSPLREVQVNLIDF 447  
Qy 188 RECOQPHYVSEVTKMLCAADPQWKTSCQDGGPLVCSLQGRMTLTGIVSWGRGAL 247  
Db 448 KKCNLDVLYDSYLTTPMMCAADLHGGRDSCQDGGPLVCEQNNRWYLAGVTSWGTGCGQ 507  
Qy 248 KDKPGVYTRVSHFLPWIRSHKTEE 271  
Db 508 RNKPGVYTKVTEVLPWIYSKMESE 531

RESULT 28  
Q86YM4 PRELIMINARY; PRT; 558 AA.  
AC Q86YM4; 01-JUN-2003 (TRENBLrel. 24, Created)  
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE Transmembrane protease serine 6.  
GN TMPRSS6.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Park T.J., Park W.J.;  
RT "Homo sapiens transmembrane protease, serine 6 (TMPRSS6) mRNA."  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY190317; AAC38062.1; -;  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0005044; F:scavenger receptor activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR InterPro; IPR001190; Srcr\_receptor.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00202; SR; 1.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS02087; SRCR 2; 1.  
DR PROSITE; PS02040; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Protease.  
SQ SEQUENCE 558 AA; 60432 MW; 4AC817FCD70D7017 CRC64;

Query Match 27.7%; Score 418; DB 4; Length 558;  
Best Local Similarity 36.5%; Pred. No. 6.9e-33;  
Matches 97; Conservative 39; Mismatches 106; Indels 24; Gaps 8;  
Qy 13 CGQKTLRPFKIIGGFTTIENQPPAAIYRRHGGSVTVVCGGSLISPCWVISAHCIFI 72  
Db 312 CGLRAMTGR--IVGGALASDKWPQVSL---HFG--TTHICGGTLIDAQWVLTAAHCFF 364  
Qy 73 DYPKK--EDYIVYLGSRSLNSNTQGMKFEVENLIHKDYSADTLAHHNDIALKIRSK 130  
Db 365 VTREKVLGKWKVAGTNSLHQLPEAA---SIAIINSNYTDE--EDDYDIALM----- 413  
Qy 131 GRCAQP---SRTIQTCLPSMYNDPQGTSCETITGKGNSTDYLPQKMTVVKLISH 187  
Db 414 -RLSKPLTSLSAHIHPACLPFMHGGTFLSLNETCMTWITFGKTRTDDTKTSPLREVQVNLIDF 472  
Qy 188 RECOQPHYVSEVTKMLCAADPQWKTSCQDGGPLVCSLQGRMTLTGIVSWGRGAL 247  
Db 473 KKCNLDVLYDSYLTTPMMCAADLHGGRDSCQDGGPLVCEQNNRWYLAGVTSWGTGCGQ 532  
Qy 248 KDKPGVYTRVSHFLPWIRSHKTEE 273

Query Match	27.7%	Score 417;	DB 4;	Length 560;
Best Local Similarity	36.2%;	Pred. No.	8.7e-33;	
Matches 104;	Conservative	40;	Mismatches	113;
			Indels	30;
Gaps	11;			
Qy	13	COOKTLRPFKIIGGEFTTIENQPWFALAIYRRHGGSVTVVCGSGLISPCWVGISATHCFI	72	
Dy	312	CGLRAMTGR--IVGGLASDSKWPQVSL---HFG--TTHICGTLDIAQWLTAHAHGF	364	

QY 73 DYPKK--EDYIVYLGSRSLNSNTGEMKFEVENLILHKDYSADTLAHNDIALLKRSKE 130  
Db 365 VTRKVLGKVVAGTGNLHQLPEAA---SIAIINSNTYDE--EDDYDIALM-----413  
QY 131 GRCAQP---SRTIQTICLPSMYNDPQFGTSCEITGFKENSTDYLYPEQLKMTVVKLISH 187  
Db 414 -RLSKPLTLGAHIHPACLPVHGQFSLNETCWITGFKTRETDDKTSPPFLREYQVNLIDF 472  
QY 188 RECOQPHYGVSEVTTKMLCADPQWKTDSCGDSGGPLVCSLQGRMTLTGIVSWGRGCAL 247  
Db 473 KKCNDYLVYDSYLTFRWMCAGDLHGGRDSCGDSGGPLVCEQNNRWYLAGVTSWGTGCGQ 532  
QY 248 KDKFQVYTRYSHFLPWIRS 266  
Db 533 RNKSGVTVKTEVLPWIS 551  
RESULT 31  
Q8CGR6 PRELIMINARY; PRT; 276 AA.  
AC Q8CGR6;  
DT 01-MAR-2003 (Tremblrel. 23, Created)  
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Glandular kallikrein KUK13.  
GN KUK13.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22325484; PubMed=12437987;  
RA Olsson A.Y., Lundwall A.;  
RT "Organization and evolution of the glandular kallikrein locus in Mus  
musculus";  
RL Biochem. Biophys. Res. Commun. 299:305-311(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Adams M., Mural R.;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY152432; AAN78420.1; -.  
DR MGD; MGI:95292; K1K13.  
DR GO; GO:004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0004285; P:trypsin activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00020; Tryp\_Spc; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
SQ SEQUENCE 276 AA; 30360 MW; 33B189C67492FDCA CRC64;  
Query Match 27.3%; Score 411.5; DB 11; Length 276;  
Best Local Similarity 38.0%; Pred. No. 1.2e-32;  
Matches 98; Conservative 40; Mismatches 81; Indels 39; Gaps 11;  
QY 26 GGEFTIENQFWAAIYRRHGGSVTVYCGGSLISPCWISATHCFIDYPK 85  
Db 39 GGYTCLPSQPQAALLRGR-----LLCGGVLPVHPKVLTAAC-----RKDGYTVHLG 88  
QY 86 RSLNSNTGEMKFEVENLILHKDYSADT--LAHNDIALKIRSKGRCAQPSRTIQT 143  
Db 89 KHALGRVNGEQAEMVRSIPHEYPQVTPTHNHDHMLLELKSPP-----VOLSSHVRTL 144  
QY 144 -----CLPSMYNDPQFGTSCEITGFKENSTDYLYPEQLKMTVVKLISHRECOQPHY 197  
Db 145 KLSADDDCLPT-----GTCCRVSGWGTTPSQVNYPTLQCANIELRSDECRQ--VYP 195

QY 198 SEVTTKMLCADPQWKTDSCGDSGGPLVCSLQGRMTLTGIVSWGR-GCALDKDPGVYTR 256  
Db 196 GKITANMLCACTGKEGKDCGDSGGPLICN--GK--LYGIISWGDFFCGQPNRPGVYTR 251  
QY 257 VSHFLPWIRS---HTKEE 271  
Db 252 VSKYLWRWIREIRNTPEQ 269  
RESULT 32  
Q8ROP5 PRELIMINARY; PRT; 638 AA.  
AC Q8ROP5;  
DT 01-JUN-2002 (Tremblrel. 21, Created)  
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Kallikrein B, plasma 1.  
GN KUKB1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC Straussberg R.;  
RA Straussberg R.;  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC026555; AAH26555.1; -.  
DR HSSP; P00761; 1AN1.  
DR MGD; MGI:102849; K1Kb1.  
DR GO; GO:0005576; C:extracellular; IEA.  
DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004295; F:trypsin activity; IEA.  
DR GO; GO:0007596; P:blood coagulation; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR001177; Apple.  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR003014; PAN\_  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00024; PAN; 4.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00005; APPLIEDOMAIN.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00223; APPLE; 4.  
DR SMART; SM00020; Tryp\_Spc; 1.  
DR PROSITE; PS00495; APPLE; 4.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Protease; Serine protease.  
SQ SEQUENCE 638 AA; 71382 MW; CC27C93F4B57C599 CRC64;  
Query Match 27.3%; Score 411; DB 11; Length 638;  
Best Local Similarity 34.2%; Pred. No. 4.1e-32;  
Matches 90; Conservative 55; Mismatches 102; Indels 16; Gaps 6;  
QY 17 TLRPREKITGSEFTIENQFWAAIYRRHGGSVTVYCGGSLISPCWISATHCFIDYPK 76  
Db 384 TKINARIYGGTVASIGEPWQVSL--QVQLVQSQTHLCCGSIIGRWLVTAACFDGTPY 441  
QY 77 KEDYIYVYLGSRSLNSNTGEMKFEVENLILHKDYSADTLAHNDIALKIRSKGRCAQ 136  
Db 442 PDVWRIYGGILSLSSEITKETSRIKELIIHQEKVS--EGNYDIALIKQTP---LNY 495  
QY 137 SRTIQTICLPSMYNDPQFGTSCEITGFG---KENSTDYLYPEQLKMTVVKLISHRECOQ 193  
Db 496 TEFQKICILPSKADTWTIYTNCWVTCGWYTKQEGTQNI-----LOKATILPVNBECCQK- 550  
QY 194 HYVGSVTVTKMLCAADPQWKTDSCGDSGGPLVCSLQGRMTLTGIVSWGRGCALDKP 253  
Db 551 KYRDIYVINKMICAGYKEGTDACKDSDSGPLVCKHSGRWQLVGTISWEGCGARKDQPGV 610

RESULT 34  
O8CJ17



```

Query Match      26.0%; Score 391.5; DB 4; Length 261;
Best Local Similarity 38.1%; Pred. No. 1.e-30;
Matches 93; Conservative 39; Mismatches 85; Indels 27; Gaps 9

QY      23 KIIGFEFTIENQ-PWFAAIYRRHRGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDI 81
      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db       37 KIVGG-YTCEENSLPYQVSL-----NSGSHFCGGSLSIQWVVAHCY----KSRIQ 84

QY      82 VYIGRSRLNSNTQGENKFEVENLILHKDYSADTLAHNDIALLKIRSEGRCAQSPRTIQ 141
      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db       85 VALGSHNIVKLEGNQFIAAKIIRHPKYARDTL--DNDIMLIKLSPP----AVINARVS 138

QY      142 TICLPSMYNDPQFGTSCSITGFGKENSTDIYLPQLNMTVVKLISHRECCQPHYGVSEVT 201
      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db       139 TISLPTA--PPAAGTECLISGMGNTLSSGADYPDELKCLDAPVLVTAQCKAS--YPGKIT 194

QY      202 TQMLCAAPDQWTKDSQQGDSGGLPVCSLQGRMTLTGIVSWGRGCALXDRPGVYTRVSHL 261
      |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db       195 NSMFCVGFLEGGKDSQGDGSGVPVVCNGQ----LQGVVSWHGCAWKNPVYTKVINYV 250

QY      262 PWIR 265
      |||
Db       251 DWIK 254

RESULT 39
Q7z410
ID Q7Z410 PRELIMINARY; PRT. 855 AA.

```

```

RESULT 40
Q812A6
Q812A6 PRELIMINARY; PRT; 453 AA.
ID Q812A6;
AC Q812A6;
DT 01-JUN-2003 (TriEMBLrel. 24, Created)
DT 01-JUN-2003 (TriEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TriEMBLrel. 25, Last annotation update)
DE Transmembrane proteinase tmprss3.
DE
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
CN NCBI_TaxID=10090;
RX [1]
RN SEQUENCE FROM N.A.
RA Rao N.V., Rao G.N., Hoidal J.R.;
RP "Genomic Organization of Murine Transmembrane Proteinases."
RT Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF479687; AAC33581.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0005044; F:scavenger receptor activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR009003; Cys_ser_trypsin.
DR InterPro; IPR002172; LBL_receptor_A.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.

```

```
DR InterPro: IPR001190; Srcr receptor.
DR Pfam: PF00057; ldl_receptor_a; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: P00722; CHYMOTRYPSIN.
DR SMART: SM00192; LDLA_1; 1.
DR SMART: SM00202; SR; 1.
DR SMART: SM00020; TRYSP; 1.
DR PROSITE: PS01209; LDLA_1; 1.
DR PROSITE: PS00668; LDLA_2; 1.
DR PROSITE: PS0287; SRCK_2; 1.
DR PROSITE: PS0240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
SQ SEQUENCE 453 AA; 49505 MW; 1E87ECD6CB3DD894 CRC64;

Query Match      25.8%; Score 389.5; DB 11; Length 453;
Best Local Similarity 35.6%; Pred. No. 3.7e-30;
Matches 94; Conservative 43; Mismatches 102; Indels 25; Gaps 10;

QY 13 CGQKTLRPRKIIIGGFTTIENQPFAAIYRRHRGGSVTVVCGGSLISPCWVISAT 71
DB 207 CGTRIGYSR--IVGNMSSLQWPMQVSL--QFG---YHLGSGIITPLWIVTAACHV 259
QY 72 IYPKKEDYIVYVYGR-SRLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSK 130
DB 260 YDLYHPKSWTVQVGLVSLMDSPVPSHL---VEKIIYHSKYKPKRLG--NDIALMKL--- 310
QY 131 GRCAQSP---SRTIQTICLPSMYNDPQFGTSCITGFGKENSVDYLYPQLKMTVVKLISH 187
DB 311 ---SEPLTDETIQTICLPSNENPFDGKLTWSNGATEDGGDASP-VLNFAAVPLISN 366
QY 188 REQQPHYVYGEVTTMCLCAADPQWKTDSCQDSGGPLVCSLQGRMTLTGIVSWGRG 247
DB 367 KICNHRDVYGGIITSPMLCAGYLVKGGVDSQDSGGPLVCSLQGRMTLTGIVSWGRG 426
QY 248 KDKPGVYTVSHPLPWRSHSTKE 271
DB 427 VNRPGVYTRITSLFDWIHQLEND 450
```

```
RESULT 41
Q99L44 PRELIMINARY; PRT; 339 AA.
AC Q99L44;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to protease, serine, 8 (proteasein).
GN PRS8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Svj, and Swiss; TISSUE=Lung;
RA Verghese G.M., Caughey G.H.;
RT "Molecular cloning and characterization of mouse prostatic, a type I
RT membrane-associated serine protease of the gamma-tryptase/prostatic
RT gene family";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Kitamura K., Takefumi N., Kimio T.;
RT "mouse serine protease.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -/- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; BC003951; AA06320.1; -.
DR EMBL; AF378086; AA06320.1; -.
DR EMBL; AF378085; AA06319.1; -.
DR EMBL; AB038244; BA82496.1; -.
DR HSSP; P00734; IUVS.
DR MGD; MGI:1923810; Prs8.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR009003; Cys_ser_trypsin.
DR InterPro: IPR001254; Peptidase_S1.
DR InterPro: IPR001314; Peptidase_S1A.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: P00722; CHYMOTRYPSIN.
DR SMART: SM00020; TRYSP; 1.
DR PROSITE: PS0240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 339 AA; 36216 MW; BC2DE88BC057AF10 CRC64;
```

```
EMBL; AF378085; AA06319.1; -.
EMBL; AB038244; BA82496.1; -.
DR HSSP; P00734; IUVS.
DR MGD; MGI:1923810; Prs8.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR009003; Cys_ser_trypsin.
DR InterPro: IPR001254; Peptidase_S1.
DR InterPro: IPR001314; Peptidase_S1A.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: P00722; CHYMOTRYPSIN.
DR SMART: SM00020; TRYSP; 1.
DR PROSITE: PS0240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 339 AA; 36216 MW; BC2DE88BC057AF10 CRC64;

Query Match      25.5%; Score 384.5; DB 11; Length 339;
Best Local Similarity 35.9%; Pred. No. 8e-30;
Matches 98; Conservative 30; Mismatches 108; Indels 37; Gaps 9

QY 13 CGQKTLRPRKIIIGGFTTIENQPFAAIYRRHRGGSVTVVCGGSLISPCWVISAT 68
DB 37 CG-AVIQPR--ITGGSAKEGQWPQV-----SITVDGNHVCGLSVSNKVVSA 84
QY 69 HCFIDYPKEDYIVYVYGR-SRLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRS 128
DB 85 HCFPREHSRAYEVKLGAGHOLDYSNDTVVHTVAQIITHSYREE--GSGDIALHLS 142
QY 129 KEGRCAQPSRTIQTICLPSMYNDPQFGTSCITGFGK-ENSTDYLYPEQLKMTVVKLISH 187
DB 143 P-----VTFSYRIPICLPAANASFPNGLHCTVTGWHVAPSVLSQTPRPLQQLVPLISR 198
QY 188 REC-----QQPHYVYGEVTTMCLCAADPQWKTDSCQDSGGPLVCSLQGRMTLTG 237
DB 199 ETGCLYNINAVPEEPH-----TIQDMLCAGYKVGKDACQDSGGPLSCPMGIWYLAG 254
QY 238 IVSWGRGKALKRPGVYTVSHPLPWRSHSTKE 270
DB 255 IVSWGDACGAPNRPVYTLTSTVASMTHHVAE 287

RESULT 42
Q8BJV6 PRELIMINARY; PRT; 340 AA.
ID Q8BJV6
AC Q8BJV6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Prostaticin.
GN PRS8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Eye;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK078696; BAC37362.1; -.
DR MGD; MGI:1923810; Prs8.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR009003; Cys_ser_trypsin.
```





RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyoko-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.,  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Stomach;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 DR ENBL; AK008927; BAB25971.1; -;  
 DR ENBL; AK003079; BAB25531.1; -;  
 DR ENBL; AK007765; BAB25241.1; -;  
 DR ENBL; AK007815; BAB25280.1; -;  
 DR ENBL; AK008729; BAB25861.1; -;  
 DR ENBL; AK008888; BAB25954.1; -;  
 DR ENBL; AK008644; BAB25226.1; -;  
 DR HSP; P00765; IGCT.  
 DR MEROPS; S01.152; -;  
 DR MGD; MGI:191373; 22000808D09Rik.  
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
 DR GO; GO:0008233; F:peptidase activity; IEA.  
 DR GO; GO:0004295; F:trypsin activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR009003; Cys Ser trypsin.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR Pfam; PF00089; trypsin; 1.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS00495; APLE; 3.  
 DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Hydrolase; Protease; Serine protease.  
 SQ SEQUENCE 263 AA; 27822 MW; 28C4487AF1A26B27 CRC64;

Query Match 25.3%; Score 382; DB 11; Length 624;  
 Best Local Similarity 33.2%; Pred. No. 1e-29;  
 Matches 86; Conservative 48; Mismatches 103; Indels 22; Gaps 7;  
 Qy 11 FCGQKTLRPF---KIIGETTTTNPWFAYIRRHGGSVTVVCGSLSPCWIS 66  
 Db 17 FCGVPAIQVLTGLSRVNGEDAIFGSNFWQVSLQDR---TGFFHFCGSLISENNVYT 72  
 Qy 67 ATFCFIDYPKEDYIVYLGSRSLNSNTGEMKPEVENLIILHKDYSADTLAHNDIALKLI 126  
 Db 73 AAHCGV---KTDVVVAGFDQGSDEENVQVUKIAQVFNKPNFSTV--RNDITLLKL 126  
 Qy 127 RSKEGRCAQSRITQITCLPSMNDPQFGTSCITGFKENSTDYLPQLKWTVVYKLIS 186  
 Db 127 ATP---AQSSEIVSAVCLPTVDDPAGTLCATTGKTKYNALKTDPKLOQAALPIVS 182  
 Qy 187 HRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGGLVCSLQGRMTLTGVSWGRGCA 246  
 Db 183 EAKKES--WGSKITDVIWICAGAS--GVSSCMGDSGGPLVCQKQDGVNTLAGIVSWGSGFC 238  
 Qy 247 LKDKPGVYTVRSHPLPWIR 265  
 Db 239 STSTPAVYARVATLMPWVQ 257

RESULT 45  
 Q9DAT3  
 ID Q9DAT3 PRELIMINARY; PRT; 624 AA.

Q9DAT3;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
 DE 1600027G01Rik protein (RIKEN CDNA 1600027G01 gene).  
 GN Fl1 OR 1600027G01Rik.  
 OS Mus musculus (Mouse).  
 OC Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Placenta;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Akawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyoko-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RT Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Strausberg R.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 DR ENBL; AK005546; BAB24114.1; -;  
 DR ENBL; BC019485; AAH19485.1; -;  
 DR HSP; P00750; IRTF.  
 DR MEROPS; S01.213; -;  
 DR MGD; MGI:99481; Fl1.  
 DR GO; GO:0005576; C:extracellular; IEA.  
 DR GO; GO:0004263; F:chymotrypsin activity; IEA.  
 DR GO; GO:0008233; F:peptidase activity; IEA.  
 DR GO; GO:0004295; F:trypsin activity; IEA.  
 DR GO; GO:0007536; P:blood coagulation; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR000177; Apple.  
 DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
 DR InterPro; IPR003014; PAN.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR Pfam; PF00024; PAN; 4.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00005; APPLEDOMAIN.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00223; APPLE; 4.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS00495; APLE; 3.  
 DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Hydrolase; Protease; Serine protease.  
 SQ SEQUENCE 624 AA; 69788 MW; 0BEDDEBC56009E97 CRC64;

Query Match 25.3%; Score 382; DB 11; Length 624;  
 Best Local Similarity 36.1%; Pred. No. 3.1e-29;  
 Matches 96; Conservative 39; Mismatches 97; Indels 34; Gaps 11

```

QY 13 CGOKTLRPRFKLIIGFEFTTIENQPMFAALYRRHGGSVTVCGSLISPCWVISATHCF 71
DB 382 CWTK-INPR--VVGAAASVGEWPMQVLIH-----SQHLICGSGIIGNQWILTAHAFCS 433
QY 72 -IDYPKEDIYVILGRSLNNTGEMKFVENILHKDYSDTLAHNDIALKIRKE 130
DB 434 GIETPKK--LRVYGGTVNGSEINEGTAFFRQEMIHDQYT--TAESGYDIALKLKLES-- 487
QY 131 GRCAQPSRTIO-TICLPSMYNDPQFGTSCETITGPF-----GKENSTDYLYPEQLKMTVVK 183
DB 488 ---AMNYTDFOPICLPSKGRNVAHTECWVTGWYTAIRGEVQST-----LQKAKVP 537
QY 184 LISHRECOQPHYVSEVTTMLCAADPQWKTSQCGDSGGLVCSLQGRMTLTGIVSWGR 243
DB 538 LVSNEEC-QTVYRRHKITNKMICAGYKEGKDKTCKGDSGGLPSCKYNGVWHLVIGTISWGE 596
QY 244 GCALKDKPGVYTVSHFLPWIRSHTK 269
DB 597 GCGQKRPVGYTNVAKYVDWILEKTQ 622

RESULT 46
QOQYN4 PRELIMINARY; PRT; 249 AA.
AC QOQYN4;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hippostasin.
GN PRS20.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Yamaguchi N., Mitsui S.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Mitsui S., Yamaguchi N.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP "cDNA cloning of a novel brain serine protease, Hippostasin.";
RL EMBL; AB016226; BAA88825.1;
DR HSP; P00763; IDPO.
DR MEROPS; S01.257; -.
DR GO; GO:0005576; C:extracellular; IDA.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR01254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00089; trypsin_1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
SQ SEQUENCE 249 AA; 27604 MW; F9FF9CB457D727D5 CRC64;

Query Match 25.3%; Score 381; DB 11; Length 249;
Best Local Similarity 35.7%; Pred. No. 1.2e-29;
Matches 90; Conservative 41; Mismatches 85; Indels 36; Gaps 10;

QY 25 IGGFEFTTIE-----NQPMFAALYRRHGGSVTVCGSLISPCWVISATHCFIDYPKK 77
DB 15 VGGETRIKGYCEPHSQPQWQVALFQKTR-----LLCGATLIAPKWLITAHHC-----RK 64
QY 78 EDYIVYILGRSLNNTGEMKFVENILHKDY--SADTLAHNDIALKIRKE--GRC 133
DB 65 PHYVILGELHLEKTDGCEORRNATESFPFHPDNNSLPNKDHRNDIMLVKNSPVFFETRA 124

```

```

QY 134 AOPSRTIOTICLPSMYNDPQFGTSCETITGFGKENSTDYLYPEQLKMTVVKLISHRECOOP 193
DB 125 VQP-LTSLPHCVAA-----GTSCILSGWITSSPQLRPLPSLRCAVNSIIIEHKECEKA 176
QY 194 HYTGSEVTTMLCAADPQWKTSQCGDSGGLVCSLQGRMTLTGIVSWGRG-CALKDKPFG 252
DB 177 --YFGNITDMLCASVRKEGKDSQCGDSGGLVPCN-----GSLQGIISWGQDPCAVTRKEG 230
QY 253 VYTRVSHFLPW 264
DB 231 VYTKVCKYFNWI 242

RESULT 47
QOQYN3 PRELIMINARY; PRT; 276 AA.
AC QOQYN3;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hippostasin prostate type (Adult male tongue cDNA, RIKEN full-length
DE enriched library, clone:2310015108 product:protease, serine, 20, full
DE insert sequence) (Adult male tongue cDNA, RIKEN full-length enriched
DE library, clone:2310040F07 product:protease, serine, 20, full insert
DE sequence).
GN PRS20.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Yamaguchi N., Mitsui S.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RX MEDLINE=20525460; PubMed=11072088;
RA Mitsui S., Okui A., Kominami K., Uemura H., Yamaguchi N.;
RT "cDNA cloning and tissue-specific splicing variants of mouse
RT hippostasin/PLSP (PRS20).";
RL Biochim. Biophys. Acta 1494:206-210 (2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Nomura R., Onno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RA The PANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RA RIKEN PANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
RN [6]

```

```

RESULT 48
Q9DGR2 PRELIMINARY; PRT; 767 AA.
AAC Q9DGR2;
ICD Q9DGR2;
AD 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Embryonic serine protease-2.
XEST-2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
NCBI_TaxID=8355;
[1]
SEQUENCE FROM N.A.
MEDLINE=20363741; PubMed=10903452;
Yanada K., Takabatake T., Takeshima K.;
"Isolation and characterization of three novel serine protease genes
from Xenopus laevis.";
RT from Xenopus laevis.;
RL Gene 252:209-216(2000).
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
EMBL; AB038437; BAB08217.1; -.
XSSP; P00766; ICHG.
DR MEROPS; S01.049; -.
DR GO; GO:0004263; F:chymotrypsin activity; IEA.
DR GO; GO:0008223; F:peptidase activity; IEA.
DR GO; GO:0004235; F:trypsin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR003003; Cys Ser trypsin.
DR InterPro; IPR02172; LDL_receptor_A.
DR InterPro; IPR001254; peptidase S1.
DR InterPro; IPR001314; peptidase S1A.
DR Pfam; PF00057; ldl_recept_a; 4.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00192; LDLa; 8.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS01209; LDLRA_1; 8.
DR PROSITE; PS00068; LDLRA_2; 2.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Hydrolase; Protease; Serine protease.
SQ SEQUENCE 767 AA; 86001 MW; B0566A38796DE96E CRC64;

Query Match 25.3%; Score 381; DB 13; Length 767;
Best Local Similarity 34.8%; Pred.No.5.le-29;
Matches 92; Conservative 38; Mismatches 108; Indels 26; Gaps 8

QY 13 CQKTLRPRFKXIGGEFTTENOPWFAAIRRHRRGGSVTVV----CGGSLISPCWVISAT 68
DB 520 CQVNSNLSVSRVIGSTFANIGNWQV-----NLQYITGLCGSIIIPKMWITAA 570

QY 69 HC-FIDYPKKBDYIVYIGRRLNSNTQGENKFFVENLIHKDYASDTLAHNDIALLKIR 127
DB 571 HCVIGSYSSASGWRVFAGTLTKPSYNNASAYP-VERIIVHPGYKSYT--VDNDIALMKLR 627

QY 128 SKRGCAQFSRTIOTICLPNMINDPFGTSCETIGFGK--ENSTDYLYPEQLKMTWVKLI 185
DB 628 DS-----ITFGYTTQVCLPSNGMFWEGATTWISGNGSYEGGSVSTY---LQYAAIPLI 680

QY 186 SHRECOQPHYYSVTTVMICAADPQWKTDSCGDSGGPLVCSLQGRMTLTGIVSWGRC 245
DB 681 DSNVCNQSYVNGQITSSMICAGYLSGGVDTCQDSGGPLVKNRNGTWLVGDTSWGDC 740

QY 245 ALKDKPGVYTVRSHFLPWIRSHTK 269
DB 741 ARANKEPGVIGNVTTFLEWISQMR 764

```

```

SEQUENCE FROM N.A.
TISSUE=Pancreas;
Strausberg R.;
Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
EMBL; BC039716; AAH39716.1; -.
GO; GO:0004263; F-tyrosinase activity; IEA.
GO; GO:0004295; F-tyrosinase activity; IEA.
GO; GO:0006508; Proteolysis and peptidolysis; IEA.
InterPro; IPR009003; Cys Ser trypsin.
InterPro; IPR001254; Peptidase S1.
InterPro; IPR001314; Peptidase_S1A.
Pfam; PF00089; trypsin; 1.
PRINTS; PR00722; CHYMOTRYPSIN.
SMART; SM00020; Tryp_Src; 1.
PROSITE; PS00240; TRYPSIN_DOM; 1.
PROSITE; PS00134; TRYPSIN_HIS; 1.
PROSITE; PS00135; TRYPSIN_SER; 1.
NON TER 1
SEQUENCE 269 AA; 28475 MW; D94C4A8E2E4BE31 CRC64;

Query Match 25.1%; Score 378.5; DB 4; Length 269;
Best Local Similarity 35.0%; Pred. No. 2.4e-29;
Matches 90; Conservative 43; Mismatches 101; Indels 23; Gaps 10

QY 13 CGQKTLPRF----KIIGFETIENQNPFAIYRHRGSGVYVCGSLISPCWVISAT 68
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
24 CGIPAIKPAISFQSIQVNGENAVLGSWPQVSL-----QDSGFHFCGGSLISQSWVVTAA 79
QY 69 HCFTDYPKEDYIVLGRSLNNGTQGMKFVENILHKDYSADTAAHNDIALLKIRS 128
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
80 HCNVY-PCR--HFVVLGEYDRSSNABPLQVLSVSRATTHPSMNSTM--NNDVTLKLAS 134
QY 129 KEGCAQPSRTIQICLPSMYNDFPGTSCEITGFGKENSVDYLPQKMTVTKLISHR 188
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
135 P-----AQYTRISPVCLASNEALTEGLTCVTTWGRLSGVGNVTFAHLQVALPLVTN 190
QY 189 ECQOPHYGSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVSWG-RCGL 247
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
191 QCRQ--YWGSSITDSMICAGA--GASSCQDGGPLVCGKNTWVLLGIVSWGKCNV 246
QY 248 KDKCVYTRVSHLPWI 264
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
247 R-APAVYTRVSKFTWI 262
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Search completed: May 25, 2004, 14:57:14
Job time : 48.7178 secs

```

OM protein - protein search, using sw model

Run on: May 25, 2004, 14:43:00 ; Search time 67.4882 Seconds  
(without alignments)  
1155.508 Million cell updates/sec

Title: US-09-880-503-5  
Perfect score: 1508  
Sequence: 1 KPSSPPBLKFCQKQKTLRP.....VSHFLPWIRSHTEKENGAL 276

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : A\_Geneseq\_29Jan04:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001s:\*
- 5: Geneseq2002s:\*
- 6: Geneseq2003as:\*
- 7: Geneseq2003bs:\*
- 8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1508	100.0	276	5	AAE16546	Aae16546 Human LMW
2	1508	100.0	323	5	AAE16548	Aae16548 Human uPA
3	1508	100.0	365	2	AAR68854	Aar68854 Delta 1-4
4	1508	100.0	386	2	AAR66266	Aar66266 Bifunctio
5	1508	100.0	390	2	AAR66245	Aar66245 Bifunctio
6	1508	100.0	390	2	AAR66247	Aar66247 Bifunctio
7	1508	100.0	392	2	AAR66260	Aar66260 Bifunctio
8	1508	100.0	392	2	AAR66264	Aar66264 Bifunctio
9	1508	100.0	392	2	AAR66255	Aar66255 Bifunctio
10	1508	100.0	392	2	AAR66259	Aar66259 Bifunctio
11	1508	100.0	392	2	AAR66261	Aar66261 Bifunctio
12	1508	100.0	392	2	AAR66258	Aar66258 Bifunctio
13	1508	100.0	392	2	AAR66263	Aar66263 Bifunctio
14	1508	100.0	392	2	AAR66254	Aar66254 Bifunctio
15	1508	100.0	392	2	AAR66256	Aar66256 Bifunctio
16	1508	100.0	392	2	AAR66257	Aar66257 Bifunctio
17	1508	100.0	393	2	AAR66251	Aar66251 Bifunctio
18	1508	100.0	393	2	AAR66253	Aar66253 Bifunctio
19	1508	100.0	393	2	AAR66249	Aar66249 Bifunctio
20	1508	100.0	393	2	AAR66252	Aar66252 Bifunctio
21	1508	100.0	393	2	AAR66244	Aar66244 Bifunctio
22	1508	100.0	393	2	AAR66250	Aar66250 Bifunctio
23	1508	100.0	393	2	AAR99597	Aar99597 Chimeric
24	1508	100.0	393	2	AAR99596	Aar99596 Chimeric
25	1508	100.0	393	2	AAR99885	Aar99885 M36: fibr

26	1508	100.0	395	2	AAR66265	Aar66265 Bifunctio
27	1508	100.0	395	2	AAR66262	Aar66262 Bifunctio
28	1508	100.0	396	2	AAR66246	Aar66246 Bifunctio
29	1508	100.0	397	2	AAR66248	Aar66248 Bifunctio
30	1508	100.0	411	1	AAP50871	Aap50871 Sequence
31	1508	100.0	411	1	AAP91684	Aap91684 Sequence
32	1508	100.0	411	1	AAP96146	Aap96146 Sequence
33	1508	100.0	411	1	AAP93589	Aap93589 Amino aci
34	1508	100.0	411	1	AAR06244	Aar06244 Urokinase
35	1508	100.0	411	2	AAR07904	Aar07904 Human pro
36	1508	100.0	411	2	AAR07902	Aar07902 Human pro
37	1508	100.0	411	2	AAR07903	Aar07903 Human pro
38	1508	100.0	411	2	AAR34584	Aar34584 Mutant hu
39	1508	100.0	411	2	AAR62991	Aar62991 Pro-uroki
40	1508	100.0	411	2	AAV39343	Aay39343 Human pro
41	1508	100.0	411	2	AAV42284	Aay42284 Human pro
42	1508	100.0	411	3	AAV92836	Aay92836 Urokinase
43	1508	100.0	411	4	AAB20489	Aab20489 Human pro
44	1508	100.0	411	4	AAB74797	Aab74797 Prourokin
45	1508	100.0	411	5	AAE16544	Aae16544 Human uro
46	1508	100.0	411	7	ADE85977	Ade85977 Human sin
47	1508	100.0	421	2	AAR10172	Aar10172 Human pro
48	1508	100.0	431	1	AAP50114	Aap50114 Sequence
49	1508	100.0	431	1	AAP60783	Aap60783 Human uro
50	1508	100.0	431	1	AAP60674	Aap60674 Modified
51	1508	100.0	431	1	AAP70238	Aap70238 Sequence
52	1508	100.0	431	1	AAP70250	Aap70250 Sequence
53	1508	100.0	431	1	AAP80430	Aap80430 Deduced A
54	1508	100.0	431	1	AAP81204	Aap81204 Pro-uroki
55	1508	100.0	431	1	AAP92119	Aap92119 Natural h
56	1508	100.0	431	1	AAR07112	Aar07112 Human pro
57	1508	100.0	431	2	AAR04253	Aar04253 Human pro
58	1508	100.0	431	2	AAR63141	Aar63141 Full leng
59	1508	100.0	431	2	AAR47903	Aar47903 Pro-uroki
60	1508	100.0	431	3	AAV50869	Aay50869 Human uro
61	1508	100.0	431	3	AAV99591	Aay99591 Human pla
62	1508	100.0	431	4	AAB84605	Aab84605 Amino aci
63	1508	100.0	431	5	AAE17128	Aae17128 Human uPA
64	1508	100.0	431	5	AAU99232	Aau99232 Human pla
65	1508	100.0	431	5	AAU99228	Aau99228 Human pla
66	1508	100.0	431	5	AAU99240	Aau99240 Human pla
67	1508	100.0	431	5	AAU99236	Aau99236 Human pla
68	1508	100.0	431	5	AAV79460	Aav79460 sc-uPA. 1
69	1508	100.0	431	6	ABR55855	Abr55855 Human uro
70	1508	100.0	431	6	ABU56547	Abu56547 Lung canc
71	1508	100.0	431	6	ABU56708	Abu56708 Lung canc
72	1508	100.0	431	6	ABU11076	Abu11076 Human uro
73	1508	100.0	431	6	ABR92137	Abr92137 Human cer
74	1508	100.0	431	7	ADE25745	Ade25745 Human pro
75	1508	100.0	432	2	AAR99888	Aar99888 M43: fibr

ALIGNMENTS

RESULT 1

AAE16546  
ID AAE16546 standard; protein; 276 AA.

XX AAE16546;

XX AC AAE16546;

XX DT 09-APR-2002 (first entry)

XX DE Human LMW-urokinase-type plasminogen activator (uPA) protein.

XX KW Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;  
stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;  
microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;  
tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;  
clotting disorder; uterine contraction disorder; respiratory disease;  
male impotence; adult respiratory distress syndrome; LMW-uPA.  
XX KW Homo sapiens.  
XX OS

```

XX PN WO200197752-A2.
XX KW Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
XX KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
XX KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
XX KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
XX KW clotting disorder; uterine contraction disorder; respiratory disease;
XX KW male impotence; adult respiratory distress syndrome; tcuPA; scuPA;
XX KW two chain urokinase; single chain urokinase; mutant; mutein.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN WO200197752-A2.
XX PD 27-DEC-2001.
XX XX
XX PF 13-JUN-2001; 2001WO-US018976.
XX XX
XX PR 20-JUN-2000; 2000US-0212874P.
XX XX
XX PA (UYPE-) UNIV PENNSYLVANIA.
XX PI Cines DB, Higazi AA;
XX PI MPI; 2002-122240/16.
XX DR N-PSDB; AAD27079.
XX DR
XX FT Composition for modulating muscle cell and tissue contractility for
XX FT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
XX FT comprising domains from urokinase-type plasminogen activator.
XX PS Claim 5; Fig 1B; 117pp; English.
XX XX
XX CC The invention relates to a composition comprising one or more domains of
XX CC urokinase-type plasminogen activator (uPA). The composition is used to
XX CC modulate the contractility and angiogenic activity of a mammalian muscle,
XX CC endothelial cell or tissue. The composition is used for treating stroke,
XX CC hypotension, thrombotic microangiopathies, heart attack, microvascular
XX CC occlusions, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
XX CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
XX CC diabetic retinopathy, wound healing, clotting disorder, uterine
XX CC contraction disorder, male impotence, respiratory disease or condition
XX CC such as asthma, adult respiratory distress syndrome, primary pulmonary
XX CC hypertension, microvascular thrombotic occlusion, and a disorder
XX CC associated with chronic intrapulmonary fibrin formation. The present
XX CC sequence is human LMW-urokinase-type plasminogen activator (uPA) protein
XX CC
XX SQ Sequence 276 AA;
XX
XX Query Match 100.0%; Score 1508; DB 5; Length 276;
XX Best Local Similarity 100.0%; Pred. No. 1.8e-125;
XX Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 KPSSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPFPAIYRRHGGSVTVVCGGSLIS 60
XX DB 1 KPSSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPFPAIYRRHGGSVTVVCGGSLIS 60
XX QY 61 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENILHKDYSADTLAHND 120
XX DB 61 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENILHKDYSADTLAHND 120
XX QY 121 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPFGTSCITGFGKENSVDLYPQLKMT 180
XX DB 121 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPFGTSCITGFGKENSVDLYPQLKMT 180
XX QY 181 VKLISHRECOQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
XX DB 181 VKLISHRECOQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
XX QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTEENGLAL 276
XX DB 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTEENGLAL 276
XX
XX RESULT 2
XX ID AAE16548
XX AC AAE16548 standard; protein; 323 AA.
XX NC AAE16548;
XX XX
XX DT 09-APR-2002 (first entry)
XX DE Human uPA delatkringle-scuPA and delatkringle-tcuPA mutant.

```

```

XX KW Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
XX KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
XX KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
XX KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
XX KW clotting disorder; uterine contraction disorder; respiratory disease;
XX KW male impotence; adult respiratory distress syndrome; tcuPA; scuPA;
XX KW two chain urokinase; single chain urokinase; mutant; mutein.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN WO200197752-A2.
XX PD 27-DEC-2001.
XX XX
XX PF 13-JUN-2001; 2001WO-US018976.
XX XX
XX PR 20-JUN-2000; 2000US-0212874P.
XX XX
XX PA (UYPE-) UNIV PENNSYLVANIA.
XX PI Cines DB, Higazi AA;
XX PI MPI; 2002-122240/16.
XX DR N-PSDB; AAD27081.
XX DR
XX FT Composition for modulating muscle cell and tissue contractility for
XX FT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
XX FT comprising domains from urokinase-type plasminogen activator.
XX PS Claim 23; Fig 1G; 117pp; English.
XX XX
XX CC The invention relates to a composition comprising one or more domains of
XX CC urokinase-type plasminogen activator (uPA). The composition is used to
XX CC modulate the contractility and angiogenic activity of a mammalian muscle,
XX CC endothelial cell or tissue. The composition is used for treating stroke,
XX CC hypotension, thrombotic microangiopathies, heart attack, microvascular
XX CC occlusions, thrombotic microangiopathies, surgically induced thrombotic
XX CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
XX CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
XX CC diabetic retinopathy, wound healing, clotting disorder, uterine
XX CC contraction disorder, male impotence, respiratory disease or condition
XX CC such as asthma, adult respiratory distress syndrome, primary pulmonary
XX CC hypertension, microvascular thrombotic occlusion, and a disorder
XX CC associated with chronic intrapulmonary fibrin formation. The present
XX CC sequence is human urokinase-type plasminogen activator (uPA) delatkringle
XX CC single chain urokinase (scuPA) and delatkringle-two chain urokinase
XX CC (tcuPA) deletion mutant
XX SQ Sequence 323 AA;
XX
XX Query Match 100.0%; Score 1508; DB 5; Length 323;
XX Best Local Similarity 100.0%; Pred. No. 2.1e-125;
XX Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0
XX
XX QY 1 KPSSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPFPAIYRRHGGSVTVVCGGSLIS 60
XX DB 48 KPSSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPFPAIYRRHGGSVTVVCGGSLIS 107
XX QY 61 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENILHKDYSADTLAHND 120
XX DB 108 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENILHKDYSADTLAHND 167
XX QY 121 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPFGTSCITGFGKENSVDLYPQLKMT 180
XX DB 168 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPFGTSCITGFGKENSVDLYPQLKMT 227
XX QY 181 VKLISHRECOQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
XX DB 228 VKLISHRECOQPHYGYSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 287
XX QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTEENGLAL 276

```

[illegible]



CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
CC to correct PN field.)  
XX  
SQ Sequence 386 AA;

Query Match 100.0%; Score 1508; DB 2; Length 386;  
Best Local Similarity 100.0%; Pred. No. 2.6e-125;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPEELKFCGQKTLRPFKLIIGGFTTIENQWPFAAIYRRHGGSVTVYVCGSLIS 60  
DB 90 KPSSPEELKFCGQKTLRPFKLIIGGFTTIENQWPFAAIYRRHGGSVTVYVCGSLIS 149

QY 61 PCWVISATHCFIDYPKKEDIYVYGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120  
DB 150 PCWVISATHCFIDYPKKEDIYVYGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 209

QY 121 IALLKIRSKGRCAPSRITQICLPSMYNDPQFGTSCIEITFGKENSTDLYLPEQLKWT 180  
DB 210 IALLKIRSKGRCAPSRITQICLPSMYNDPQFGTSCIEITFGKENSTDLYLPEQLKWT 269

QY 181 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240  
DB 270 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 329

QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 276  
DB 330 WGRGCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 365

RESULT 5  
AAR66245  
ID AAR66245 standard; protein; 390 AA.  
XX  
AC AAR66245;  
XX  
XX 25-MAR-2003 (revised)  
DT 17-AUG-1995 (first entry)  
XX  
DE Bifunctional urokinase variant M12.  
XX  
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
KW urokinase; variant; mutein.  
XX  
OS Synthetic.  
XX  
XX Key Location/Qualifiers  
FH Region 1..365  
FT /label= M4  
FT /note= "unglycosylated prourokinase (Ser47-Leu411)"  
FT Disulfide-bond 4..85  
FT Disulfide-bond 25..67  
FT Disulfide-bond 56..80  
FT Disulfide-bond 102..233  
FT Disulfide-bond 143..159  
FT Disulfide-bond 181..222  
FT Disulfide-bond 247..316  
FT Disulfide-bond 279..295  
FT Disulfide-bond 306..334  
FT Disulfide-bond 366..371  
FT Region /label= X1  
FT Region 372..390  
FT /label= Y1  
XX  
XX DE4323754-Cl.  
XX  
XX 01-DEC-1994.  
XX  
XX 15-JUL-1993; 93DE-04323754.  
XX  
XX 15-JUL-1993; 93DE-04323754.  
XX  
XX (CHEF ) GRUENENTHAL GMBH.  
FA

XX Steffens GU, Wnendt S, Schneider J, Heinz-Wieland R;  
PI Saunders DJ;  
XX  
XX WPI; 1995-015191/03.  
XX  
XX New bifunctional urokinase derivs and related plasmids - with improved  
PT fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
PT cerebral infarct, pulmonary embolism, etc.  
XX  
XX Example 1; Page 10 and Fig 1; 34pp; German.  
XX  
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
CC are claimed (see features table). Sequences AAR66244-R66266 are specific  
CC examples of such derivs, which have both improved fibrinolytic and  
CC thrombin-inhibiting activities, compared to known plasminogen activators  
CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
CC to correct PN field.)  
XX  
XX  
SQ Sequence 390 AA;

Query Match 100.0%; Score 1508; DB 2; Length 390;  
Best Local Similarity 100.0%; Pred. No. 2.7e-125;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPEELKFCGQKTLRPFKLIIGGFTTIENQWPFAAIYRRHGGSVTVYVCGSLIS 60  
DB 90 KPSSPEELKFCGQKTLRPFKLIIGGFTTIENQWPFAAIYRRHGGSVTVYVCGSLIS 149

QY 61 PCWVISATHCFIDYPKKEDIYVYGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120  
DB 150 PCWVISATHCFIDYPKKEDIYVYGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 209

QY 121 IALLKIRSKGRCAPSRITQICLPSMYNDPQFGTSCIEITFGKENSTDLYLPEQLKWT 180  
DB 210 IALLKIRSKGRCAPSRITQICLPSMYNDPQFGTSCIEITFGKENSTDLYLPEQLKWT 269

QY 181 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240  
DB 270 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 329

QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 276  
DB 330 WGRGCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 365

RESULT 6  
AAR66247  
ID AAR66247 standard; protein; 390 AA.  
XX  
AC AAR66247;  
XX  
XX 25-MAR-2003 (revised)  
DT 17-AUG-1995 (first entry)  
XX  
XX Bifunctional urokinase variant M14.  
XX  
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
KW urokinase; variant; mutein.  
XX  
OS Synthetic.  
XX  
XX Key Location/Qualifiers  
FH Region 1..365  
FT /label= M4  
FT /note= "unglycosylated prourokinase (Ser47-Leu411)"  
FT Disulfide-bond 4..85  
FT Disulfide-bond 25..67  
FT Disulfide-bond 56..80  
FT Disulfide-bond 102..233  
FT Disulfide-bond 143..159  
FT

FT Disulfide-bond 151..222  
FT Disulfide-bond 247..316  
FT Disulfide-bond 279..295  
FT Disulfide-bond 306..334  
FT Region /label= X1  
FT Region /label= Y1  
XX DE4323754-C1.  
XX PD 01-DEC-1994.  
XX PF 15-JUL-1993; 93DE-04323754.  
XX PR 15-JUL-1993; 93DE-04323754.  
XX PA (CHEF ) GRUENENTHAL GMBH.  
XX PI Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
PI Saunders DJ;  
XX WPI; 1995-015191/03.  
XX New bifunctional urokinase derivs and related plasmids - with improved  
PT fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
PT cerebral infarct, pulmonary embolism, etc.  
XX Example 1; Page 10 and Fig 1; 34pp; German.  
CC Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
CC are claimed (see features table). Sequences AAR66244-R66266 are specific  
CC examples of such derivs. which have both improved fibrinolytic and  
CC thrombin-inhibiting activities, compared to known plasminogen activators  
CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
CC to correct PN field.)  
XX Sequence 390 AA;  
Query Match 100.0%; Score 1508; DB 2; Length 390;  
Best Local Similarity 100.0%; Pred. No. 2.7e-125;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KPSSPPEELKFCQGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRHRGGSVTVCGGSLIS 60  
DB 90 KPSSPPEELKFCQGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRHRGGSVTVCGGSLIS 149  
QY 61 PCWVISATCFIDYPKKEDYIVYLGSRNSNTQGMKFEVENLILHKDYADTLAHND 120  
DB 150 PCWVISATCFIDYPKKEDYIVYLGSRNSNTQGMKFEVENLILHKDYADTLAHND 209  
QY 121 IALLKIRSKGRCAPSRITQITCLPSMYNDPQFGTSCEITGFKENSTDYLYPQLKMT 180  
DB 210 IALLKIRSKGRCAPSRITQITCLPSMYNDPQFGTSCEITGFKENSTDYLYPQLKMT 269  
QY 181 VKLISHRCQQPHYGVSVTFMCAADPQKNTSCQDSCGGLVCSLQGRWTLTGVS 240  
DB 270 VKLISHRCQQPHYGVSVTFMCAADPQKNTSCQDSCGGLVCSLQGRWTLTGVS 329  
QY 241 WGRGCALKDKPGYTVRSHFLPWIRSHKTEENGLAL 276  
DB 330 WGRGCALKDKPGYTVRSHFLPWIRSHKTEENGLAL 365  
RESULT 7  
AAR66260  
ID AAR66260 standard; protein; 392 AA.  
XX  
AC AAR66260;  
XX  
DT 25-MAR-2003 (revised)

DT 17-AUG-1995 (first entry)  
XX Bifunctional urokinase variant M27.  
DE fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
KW urokinase; variant; mutein.  
XX Synthetic.  
XX Key Location/Qualifiers  
FT Region 1..365  
FT /label= M4  
FT /notes="unglycosylated prourokinase (Ser47-Leu411)"  
FT Disulfide-bond 4..85  
FT Disulfide-bond 25..67  
FT Disulfide-bond 56..80  
FT Disulfide-bond 102..233  
FT Disulfide-bond 143..159  
FT Disulfide-bond 151..222  
FT Disulfide-bond 247..316  
FT Disulfide-bond 279..295  
FT Disulfide-bond 306..334  
FT Region 366..371  
FT /label= X1  
FT Region 372..392  
FT /label= Y1  
XX DE4323754-C1.  
XX PD 01-DEC-1994.  
XX PF 15-JUL-1993; 93DE-04323754.  
XX PR 15-JUL-1993; 93DE-04323754.  
XX (CHEF ) GRUENENTHAL GMBH.  
XX Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
PI Saunders DJ;  
XX WPI; 1995-015191/03.  
XX New bifunctional urokinase derivs and related plasmids - with improved  
PT fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
PT cerebral infarct, pulmonary embolism, etc.  
XX Example 1; Page 11 and Fig 1; 34pp; German.  
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
CC are claimed (see features table). Sequences AAR66244-R66266 are specific  
CC examples of such derivs. which have both improved fibrinolytic and  
CC thrombin-inhibiting activities, compared to known plasminogen activators  
CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
CC to correct PN field.)  
XX Sequence 392 AA;  
Query Match 100.0%; Score 1508; DB 2; Length 392;  
Best Local Similarity 100.0%; Pred. No. 2.7e-125;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0  
QY 1 KPSSPPEELKFCQGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRHRGGSVTVCGGSLIS 60  
DB 90 KPSSPPEELKFCQGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRHRGGSVTVCGGSLIS 149  
QY 61 PCWVISATCFIDYPKKEDYIVYLGSRNSNTQGMKFEVENLILHKDYADTLAHND 120  
DB 150 PCWVISATCFIDYPKKEDYIVYLGSRNSNTQGMKFEVENLILHKDYADTLAHND 209  
QY 121 IALLKIRSKGRCAPSRITQITCLPSMYNDPQFGTSCEITGFKENSTDYLYPQLKMT 180

Db 210 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCITGFGKENSTDYLYPEQLKMT 269

QY 181 VVKLISHRECOQPHYHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240

Db 270 VVKLISHRECOQPHYHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 329

QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 276

Db 330 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 365

RESULT 8

AAR66264

ID AAR66264 standard; protein; 392 AA.

XX AAR66264;

XX 25-MAR-2003 (revised)

DT 17-AUG-1995 (first entry)

XX Bifunctional urokinase variant M31.

DE fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;

XX urokinase; variant; mutein.

XX Synthetic.

XX Key Location/Qualifiers

FT 1. .365

FT /label= M4

FT /note= "unglycosylated prourokinase (Ser47-Leu411)"

FT Disulfide-bond 4. .85

FT Disulfide-bond 25. .67

FT Disulfide-bond 56. .80

FT Disulfide-bond 102. .233

FT Disulfide-bond 143. .159

FT Disulfide-bond 151. .222

FT Disulfide-bond 247. .316

FT Disulfide-bond 279. .295

FT Disulfide-bond 306. .334

FT Region 366. .371

FT /label= X1

FT 372. .392

FT /label= Y1

XX DB4323754-C1.

XX 01-DEC-1994.

XX 15-JUL-1993; 93DE-04323754.

XX 15-JUL-1993; 93DE-04323754.

XX (CHEF ) GRUENTHAL GMBH.

XX Steffens GJ, Whendt S, Schneider J, Heinzel-Wieland R;

PI Saunders DJ;

XX WPI; 1995-015191/03.

XX New bifunctional urokinase derivs and related plasmids - with improved

PT fibrinolytic and thrombin inhibiting activities, for treating cardiac and

PT cerebral infarct, pulmonary embolism, etc.

XX Example 1; Page 11 and Fig 1; 34pp; German.

XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1

CC are claimed (see features table). Sequences AAR66264-R66266 are specific

CC examples of such derivs. which have both improved fibrinolytic and

CC thrombin-inhibiting activities, compared to known plasminogen activators

CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,

CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac

CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003

CC to correct PN field.)

XX Sequence 392 AA;

QY Query Match 100.0%; Score 1508; DB 2; Length 392;

Best Local Similarity 100.0%; Pred. NO. 2.7e-125; Indels 0; Gaps 0

Matches 276; Conservative 0; Mismatches 0;

QY 1 KPSSPPEELKFCQGQKTLRPRFKIIGBFTTIENQWFAAIYRRHRGGSVTVVCGSLIS 60

Db 90 KPSSPPEELKFCQGQKTLRPRFKIIGBFTTIENQWFAAIYRRHRGGSVTVVCGSLIS 149

QY 61 PCWVISATHCFIDYPKKEDIIVYLGSRSLNSNTQGMKFEVENLIHKDYSDTLAHND 120

Db 150 PCWVISATHCFIDYPKKEDIIVYLGSRSLNSNTQGMKFEVENLIHKDYSDTLAHND 209

QY 121 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCITGFGKENSTDYLYPEQLKMT 180

Db 210 IALLKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCITGFGKENSTDYLYPEQLKMT 269

QY 181 VVKLISHRECOQPHYHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240

Db 270 VVKLISHRECOQPHYHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 329

QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 276

Db 330 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 365

RESULT 9

AAR66255

ID AAR66255 standard; protein; 392 AA.

XX AAR66255;

XX 25-MAR-2003 (revised)

DT 17-AUG-1995 (first entry)

XX Bifunctional urokinase variant M22.

XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;

XX urokinase; variant; mutein.

XX Synthetic.

XX Key Location/Qualifiers

FT 1. .365

FT /label= M4

FT /note= "unglycosylated prourokinase (Ser47-Leu411)"

FT Disulfide-bond 4. .85

FT Disulfide-bond 25. .67

FT Disulfide-bond 56. .80

FT Disulfide-bond 102. .233

FT Disulfide-bond 143. .159

FT Disulfide-bond 151. .222

FT Disulfide-bond 247. .316

FT Disulfide-bond 279. .295

FT Disulfide-bond 306. .334

FT Region 366. .371

FT /label= X1

FT 372. .392

FT /label= Y1

XX DB4323754-C1.

XX 01-DEC-1994.

XX 15-JUL-1993; 93DE-04323754.

XX 15-JUL-1993; 93DE-04323754.

XX (CHEF ) GRUENTHAL GMBH.

PI Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
PI Saunders DJ;  
XX WPI; 1995-015191/03.  
XX New bifunctional urokinase derivs and related plasmids - with improved  
PT fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
PT cerebral infarct, pulmonary embolism, etc.  
XX Example 1; Page 10 and Fig 1; 34pp; German.  
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
CC are claimed (see features table). Sequences AAR66244-R66266 are specific  
CC examples of such derivs, which have both improved fibrinolytic and  
CC thrombin-inhibiting activities, compared to known plasminogen activators  
CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
CC to correct PN field.)  
XX Sequence 392 AA;

Query Match 100.0%; Score 1508; DB 2; Length 392;  
Best Local Similarity 100.0%; Pred. No. 2.7e-125;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KPSSPPEELKFCQGKTLRPRFKIIGGEFTTIENQPFWFAAIYRRHGGSVTVVCGSLIS 60  
DB 90 KPSSPPEELKFCQGKTLRPRFKIIGGEFTTIENQPFWFAAIYRRHGGSVTVVCGSLIS 149  
QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120  
DB 150 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 209  
QY 121 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSVDLYPQLKMT 180  
DB 210 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSVDLYPQLKMT 269  
QY 181 VVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240  
DB 270 VVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 329  
QY 241 WGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL 276  
DB 330 WGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL 365

RESULT 10  
AAR66259  
ID AAR66259 standard; protein; 392 AA.  
XX AAR66259;  
AC AAR66259;  
XX 25-MAR-2003 (revised)  
DT 17-AUG-1995 (first entry)  
XX Bifunctional urokinase variant M26.  
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
XX urokinase; variant; mutein.  
XX Synthetic.  
XX Key Location/Qualifiers  
FH 1..365  
FT /label= M4  
FT /note= "unglycosylated prourokinase (Ser47-Leu411)"  
FT Disulfide-bond 247..316  
FT Disulfide-bond 279..295  
FT Disulfide-bond 306..334  
FT Region 366..371  
FT /label= X1  
FT Region 372..392  
FT /label= Y1  
XX DB4323754-C1.  
XX 01-DEC-1994.  
XX 15-JUL-1993; 93DE-04323754.  
XX 15-JUL-1993; 93DE-04323754.  
XX (CHEF) GRUENENTHAL GMBH.  
XX Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
XX Saunders DJ;  
XX WPI; 1995-015191/03.  
XX New bifunctional urokinase derivs and related plasmids - with improved  
PT fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
PT cerebral infarct, pulmonary embolism, etc.  
XX Example 1; Page 11 and Fig 1; 34pp; German.  
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
CC are claimed (see features table). Sequences AAR66244-R66266 are specific  
CC examples of such derivs, which have both improved fibrinolytic and  
CC thrombin-inhibiting activities, compared to known plasminogen activators  
CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
CC to correct PN field.)  
XX Sequence 392 AA;

FT Disulfide-bond 247..316  
FT Disulfide-bond 279..295  
FT Disulfide-bond 306..334  
FT Region 366..371  
FT /label= X1  
FT Region 372..392  
FT /label= Y1  
XX DB4323754-C1.  
XX 01-DEC-1994.  
XX 15-JUL-1993; 93DE-04323754.  
XX 15-JUL-1993; 93DE-04323754.  
XX (CHEF) GRUENENTHAL GMBH.  
XX Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
XX Saunders DJ;  
XX WPI; 1995-015191/03.  
XX New bifunctional urokinase derivs and related plasmids - with improved  
PT fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
PT cerebral infarct, pulmonary embolism, etc.  
XX Example 1; Page 11 and Fig 1; 34pp; German.  
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
CC are claimed (see features table). Sequences AAR66244-R66266 are specific  
CC examples of such derivs, which have both improved fibrinolytic and  
CC thrombin-inhibiting activities, compared to known plasminogen activators  
CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
CC to correct PN field.)  
XX Sequence 392 AA;

Query Match 100.0%; Score 1508; DB 2; Length 392;  
Best Local Similarity 100.0%; Pred. No. 2.7e-125;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KPSSPPEELKFCQGKTLRPRFKIIGGEFTTIENQPFWFAAIYRRHGGSVTVVCGSLIS 60  
DB 90 KPSSPPEELKFCQGKTLRPRFKIIGGEFTTIENQPFWFAAIYRRHGGSVTVVCGSLIS 149  
QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120  
DB 150 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 209  
QY 121 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSVDLYPQLKMT 180  
DB 210 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSVDLYPQLKMT 269  
QY 181 VVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240  
DB 270 VVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 329  
QY 241 WGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL 276  
DB 330 WGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL 365

RESULT 11  
AAR66261  
ID AAR66261 standard; protein; 392 AA.  
XX AAR66261;  
AC AAR66261;  
XX 25-MAR-2003 (revised)  
DT 17-AUG-1995 (first entry)

XX Bifunctional urokinase variant M28.  
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
XX urokinase; variant; mutein.  
XX Synthetic.  
XX Key Location/Qualifiers  
XX Region 1..365  
XX /label= M4  
XX /note= "unglycosylated prourokinase (Ser47-Leu411)"  
XX Disulfide-bond 4..85  
XX Disulfide-bond 25..67  
XX Disulfide-bond 56..80  
XX Disulfide-bond 102..233  
XX Disulfide-bond 143..159  
XX Disulfide-bond 151..222  
XX Disulfide-bond 247..316  
XX Disulfide-bond 279..295  
XX Disulfide-bond 306..334  
XX Disulfide-bond 366..371  
XX Region /label= X1  
XX Region /label= Y1  
XX DE4323754-Cl.  
XX 01-DEC-1994.  
XX 15-JUL-1993; 93DE-04323754.  
XX 15-JUL-1993; 93DE-04323754.  
XX (CHEF ) GRUENENTHAL GMBH.  
XX Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
XX Saunders DJ;  
XX WPI; 1995-015191/03.  
XX New bifunctional urokinase derivs and related plasmids - with improved  
XX fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
XX cerebral infarct, pulmonary embolism, etc.  
XX Example 1; Page 11 and Fig 1; 34pp; German.  
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
XX are claimed (see features table). Sequences AAR6244-R66266 are specific  
XX examples of such derivs. which have both improved fibrinolytic and  
XX thrombin-inhibiting activities, compared to known plasminogen activators  
XX or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
XX e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
XX and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
XX to correct PN field.)  
XX SQ Sequence 392 AA;  
Query Match 100.0%; Score 1508; DB 2; Length 392;  
Best Local Similarity 100.0%; Pred No. 2.7e-125;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KPSSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQWFAAIVRRHGGSVTVVCGSLIS 60  
DB 90 KPSSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQWFAAIVRRHGGSVTVVCGSLIS 149  
QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRRLNSNTQGEWKEVENLIILHKDYSADTLAHND 120  
DB 150 PCWVISATHCFIDYPKKEDYIVYLGSRRLNSNTQGEWKEVENLIILHKDYSADTLAHND 209  
QY 121 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQGTSCBITGFGKENS'DYLYPEQLKMT 180  
DB 210 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQGTSCBITGFGKENS'DYLYPEQLKMT 269

QY 181 VVKLISHRECCQPHYVGVSEVTTKMLCAADPQWKTSQCGDSGGPLVCSLQGRMTLTGIVS 240  
DB 270 VVKLISHRECCQPHYVGVSEVTTKMLCAADPQWKTSQCGDSGGPLVCSLQGRMTLTGIVS 329  
QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276  
DB 330 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 365  
RESULT 12  
AAR66258  
ID AAR66258 standard; protein; 392 AA.  
XX AC AAR66258;  
XX 25-MAR-2003 (revised)  
XX 17-AUG-1995 (first entry)  
XX Bifunctional urokinase variant M25.  
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
XX urokinase; variant; mutein.  
XX Synthetic.  
XX Key Location/Qualifiers  
XX Region 1..365  
XX /label= M4  
XX /note= "unglycosylated prourokinase (Ser47-Leu411)"  
XX Disulfide-bond 4..85  
XX Disulfide-bond 25..67  
XX Disulfide-bond 56..80  
XX Disulfide-bond 102..233  
XX Disulfide-bond 143..159  
XX Disulfide-bond 151..222  
XX Disulfide-bond 247..316  
XX Disulfide-bond 279..295  
XX Disulfide-bond 306..334  
XX Disulfide-bond 366..371  
XX Region /label= X1  
XX Region /label= Y1  
XX DE4323754-Cl.  
XX 01-DEC-1994.  
XX 15-JUL-1993; 93DE-04323754.  
XX 15-JUL-1993; 93DE-04323754.  
XX (CHEF ) GRUENENTHAL GMBH.  
XX Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
XX Saunders DJ;  
XX WPI; 1995-015191/03.  
XX New bifunctional urokinase derivs and related plasmids - with improved  
XX fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
XX cerebral infarct, pulmonary embolism, etc.  
XX Example 1; Page 11 and Fig 1; 34pp; German.  
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
XX are claimed (see features table). Sequences AAR6244-R66266 are specific  
XX examples of such derivs. which have both improved fibrinolytic and  
XX thrombin-inhibiting activities, compared to known plasminogen activators  
XX or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
XX e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
XX and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
XX to correct PN field.)

```

XX SQ Sequence 392 AA;
Query Match 100.0%; Score 1508; DB 2; Length 392;
Best Local Similarity 100.0%; Pred. No. 2.7e-125;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQGOKTLRPFKLIIGBEFTTIENQWPFAAIYRRHGGSVTVVCGSLIS 60
DB 90 KPSSPPEELKFCQGOKTLRPFKLIIGBEFTTIENQWPFAAIYRRHGGSVTVVCGSLIS 149
QY 61 PCWISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHHND 120
DB 150 PCWISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHHND 209
QY 121 IALLKIRSKGRCAPSRITQICLPSMYNDPQFGTSCITGFGKENSTDYLYPEQLKMT 180
DB 210 IALLKIRSKGRCAPSRITQICLPSMYNDPQFGTSCITGFGKENSTDYLYPEQLKMT 269
QY 181 VVKLISHRECCQPHYVGSVTTKMLCAADPWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
DB 270 VVKLISHRECCQPHYVGSVTTKMLCAADPWKTDSCQDGGPLVCSLQGRMTLTGIVS 329
QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
DB 330 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 365

RESULT 13
AAR66263
ID AAR66263 standard; protein; 392 AA.
AC AAR66263;
XX
XX 25-MAR-2003 (revised)
DT 17-AUG-1995 (first entry)
XX
DE Bifunctional urokinase variant M30.
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase (Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Region 366..371
FT /label= X1
FT Region 372..392
FT /label= Y1
XX
XX DE4323754-Cl.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-04323754.
XX
XX 15-JUL-1993; 93DE-04323754.
XX (CHEF ) GRUENTHAL GMBH.
XX
XX Steffens GU, Wnendt S, Schneider J, Heinzel-Wieland R;

```

```

PI Saunders DJ;
XX
XX WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with improved
PT fibrinolytic and thrombin inhibiting activities, for treating cardiac and
PT cerebral infarct, pulmonary embolism, etc.
XX
XX Example 1; Page 11 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1
CC are claimed (see features table). Sequences AAR66244-R66266 are specific
CC examples of such derivs. which have both improved fibrinolytic and
CC thrombin-inhibiting activities, compared to known plasminogen activators
CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,
CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac
CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003
XX to correct PN field.)
XX
XX Sequence 392 AA;
Query Match 100.0%; Score 1508; DB 2; Length 392;
Best Local Similarity 100.0%; Pred. No. 2.7e-125;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 KPSSPPEELKFCQGOKTLRPFKLIIGBEFTTIENQWPFAAIYRRHGGSVTVVCGSLIS 60
DB 90 KPSSPPEELKFCQGOKTLRPFKLIIGBEFTTIENQWPFAAIYRRHGGSVTVVCGSLIS 149
QY 61 PCWISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHHND 120
DB 150 PCWISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHHND 209
QY 121 IALLKIRSKGRCAPSRITQICLPSMYNDPQFGTSCITGFGKENSTDYLYPEQLKMT 180
DB 210 IALLKIRSKGRCAPSRITQICLPSMYNDPQFGTSCITGFGKENSTDYLYPEQLKMT 269
QY 181 VVKLISHRECCQPHYVGSVTTKMLCAADPWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
DB 270 VVKLISHRECCQPHYVGSVTTKMLCAADPWKTDSCQDGGPLVCSLQGRMTLTGIVS 329
QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
DB 330 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 365

RESULT 14
AAR66254
ID AAR66254 standard; protein; 392 AA.
XX
XX AAR66254;
XX
XX 25-MAR-2003 (revised)
DT 17-AUG-1995 (first entry)
XX
XX Bifunctional urokinase variant M21.
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase (Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316

```

FT Disulfide-bond 279. .295  
FT Disulfide-bond 306. .334  
FT Region 366. .371  
FT /label= X1  
FT /label= Y1  
FT /label= Y1  
XX DE4323754-C1.  
XX 01-DEC-1994.  
XX 15-JUL-1993; 93DE-04323754.  
XX 15-JUL-1993; 93DE-04323754.  
XX (CHEF ) GRUENENTHAL GMBH.  
XX Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
XX Saunders DJ;  
XX WPI; 1995-015191/03.  
XX New bifunctional urokinase derivs and related plasmids - with improved  
XX fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
XX cerebral infarct, pulmonary embolism, etc.  
XX Example 1; Page 10 and Fig 1; 34pp; German.  
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
XX are claimed (see features table). Sequences AAR66244-R66266 are specific  
XX examples of such derivs. which have both improved fibrinolytic and  
XX thrombin-inhibiting activities, compared to known plasminogen activators  
XX or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
XX e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
XX and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
XX to correct PN field.)  
XX Sequence 392 AA;  
XX  
XX Query Match 100.0%; Score 1508; DB 2; Length 392;  
XX Best Local Similarity 100.0%; Pred. No. 2.7e-125; Indels 0; Gaps 0;  
XX Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KPSSPPEELKFCQGKTLPRFKIIGGEFTTIENQPWFAAIYRRHRGGSVTVVCGSLIS 60  
DB 90 KPSSPPEELKFCQGKTLPRFKIIGGEFTTIENQPWFAAIYRRHRGGSVTVVCGSLIS 149  
QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120  
DB 150 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 209  
QY 121 IALLKIRSKERCAQPSRTIQTICLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMT 180  
DB 210 IALLKIRSKERCAQPSRTIQTICLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMT 269  
QY 181 VKLISHRECOQPHYGVSEVTTKMLCAADPQWTKDSQGDSPGVCSIQGRMTLTIGVS 240  
DB 270 VKLISHRECOQPHYGVSEVTTKMLCAADPQWTKDSQGDSPGVCSIQGRMTLTIGVS 329  
QY 241 WGRGCALKDKPGYVTVSHFLPWIRSHTEENGLAL 276  
DB 330 WGRGCALKDKPGYVTVSHFLPWIRSHTEENGLAL 365  
RESULT 15  
ID AAR66256  
XX AAR66256 standard; protein; 392 AA.  
XX AC AAR66256;  
XX 25-MAR-2003 (revised)  
DT 17-AUG-1995 (first entry)  
XX

DE Bifunctional urokinase variant M23.  
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
XX urokinase; variant; mutein.  
XX Synthetic.  
XX Key Location/Qualifiers  
XX Region 1. .365  
XX /label= M4  
XX /note= "unglycosylated prourokinase(Ser47-Leu411)"  
FT Disulfide-bond 4. .85  
FT Disulfide-bond 25. .67  
FT Disulfide-bond 56. .80  
FT Disulfide-bond 102. .233  
FT Disulfide-bond 143. .159  
FT Disulfide-bond 151. .222  
FT Disulfide-bond 247. .316  
FT Disulfide-bond 279. .295  
FT Disulfide-bond 306. .334  
FT Disulfide-bond 366. .371  
FT Region /label= X1  
FT Region 372. .392  
FT /label= Y1  
XX DE4323754-C1.  
XX 01-DEC-1994.  
XX 15-JUL-1993; 93DE-04323754.  
XX 15-JUL-1993; 93DE-04323754.  
XX (CHEF ) GRUENENTHAL GMBH.  
XX Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
XX Saunders DJ;  
XX WPI; 1995-015191/03.  
XX New bifunctional urokinase derivs and related plasmids - with improved  
XX fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
XX cerebral infarct, pulmonary embolism, etc.  
XX Example 1; Page 10 and Fig 1; 34pp; German.  
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
XX are claimed (see features table). Sequences AAR66244-R66266 are specific  
XX examples of such derivs. which have both improved fibrinolytic and  
XX thrombin-inhibiting activities, compared to known plasminogen activators  
XX or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
XX e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
XX and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
XX to correct PN field.)  
XX Sequence 392 AA;  
XX  
XX Query Match 100.0%; Score 1508; DB 2; Length 392;  
XX Best Local Similarity 100.0%; Pred. No. 2.7e-125; Indels 0; Gaps 0;  
XX Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KPSSPPEELKFCQGKTLPRFKIIGGEFTTIENQPWFAAIYRRHRGGSVTVVCGSLIS 60  
DB 90 KPSSPPEELKFCQGKTLPRFKIIGGEFTTIENQPWFAAIYRRHRGGSVTVVCGSLIS 149  
QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120  
DB 150 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 209  
QY 121 IALLKIRSKERCAQPSRTIQTICLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMT 180  
DB 210 IALLKIRSKERCAQPSRTIQTICLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMT 269

QY 181 VVKLISHRECCQPHYGSEVTTKMLCAADPOWKTDSCQDGGGGLVCSLQGRMTLTGIVS 240  
 DB 270 VVKLISHRECCQPHYGSEVTTKMLCAADPOWKTDSCQDGGGGLVCSLQGRMTLTGIVS 329  
 QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGAL 276  
 DB 330 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGAL 365

RESULT 16  
 AAR66257  
 ID AAR66257 standard; protein; 392 AA.  
 XX AAR66257;  
 AC AAR66257;  
 XX  
 XX 25-MAR-2003 (revised)  
 DT 17-AUG-1995 (first entry)  
 XX  
 XX Bifunctional urokinase variant M24.  
 DE  
 XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
 KW urokinase; variant; mutein.  
 XX  
 XX Synthetic.  
 OS  
 FH Key Location/Qualifiers  
 FT Region 1..365  
 FT /label= M4  
 FT /note= "unglycosylated prourokinase (Ser47-Leu411)"  
 FT Disulfide-bond 25..67  
 FT Disulfide-bond 56..80  
 FT Disulfide-bond 102..233  
 FT Disulfide-bond 143..159  
 FT Disulfide-bond 151..222  
 FT Disulfide-bond 247..316  
 FT Disulfide-bond 279..295  
 FT Disulfide-bond 306..334  
 FT Region 366..371  
 FT /label= X1  
 FT Region 372..392  
 FT /label= Y1  
 XX  
 PN DE4323754-Cl.  
 XX  
 PD 01-DEC-1994.  
 XX  
 XX 15-JUL-1993; 93DE-04323754.  
 XX  
 PR 15-JUL-1993; 93DE-04323754.  
 XX  
 XX (CHEF ) GRUENENTHAL GMBH.  
 XX  
 PI Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
 PI Saunders DJ;  
 XX  
 DR WPI; 1995-015191/03.  
 XX  
 XX New bifunctional urokinase derivs and related plasmids - with improved  
 PT fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
 PT cerebral infarct, pulmonary embolism, etc.  
 XX  
 PS Example 1; Page 11 and Fig 1; 34pp; German.  
 XX  
 CC Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
 CC are claimed (see features table). Sequences AAR66244-R66266 are specific  
 CC examples of such derivs, which have both improved fibrinolytic and  
 CC thrombin-inhibiting activities, compared to known plasminogen activators  
 CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
 CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
 CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
 CC to correct PN field.)  
 XX

SQL Sequence 392 AA;  
 Query Match 100.0%; Score 1508; DB 2; Length 392;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-125;  
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 KPSSPPEELKFOGQKTLRPRFKIIGGEFTTIENQFWFAAIYRRHGGSVTVVCGSLIS 60  
 DB 90 KPSSPPEELKFOGQKTLRPRFKIIGGEFTTIENQFWFAAIYRRHGGSVTVVCGSLIS 149  
 QY 61 PCWVISATHCFIDYPKKEDIYVYLGSRRLNSNTQEMKFEVENLILHKDYSADTTLAHND 120  
 DB 150 PCWVISATHCFIDYPKKEDIYVYLGSRRLNSNTQEMKFEVENLILHKDYSADTTLAHND 209  
 QY 121 IALLKIRSEGRCAQPSRTIQICLPMSYNDPQGTSCETGFGKENSVDLYLPQLKMT 180  
 DB 210 IALLKIRSEGRCAQPSRTIQICLPMSYNDPQGTSCETGFGKENSVDLYLPQLKMT 269  
 QY 181 VVKLISHRECCQPHYGSEVTTKMLCAADPOWKTDSCQDGGGGLVCSLQGRMTLTGIVS 240  
 DB 270 VVKLISHRECCQPHYGSEVTTKMLCAADPOWKTDSCQDGGGGLVCSLQGRMTLTGIVS 329  
 QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGAL 276  
 DB 330 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGAL 365

RESULT 17  
 AAR66251  
 ID AAR66251 standard; protein; 393 AA.  
 XX  
 XX AAR66251;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 17-AUG-1995 (first entry)  
 XX  
 XX Bifunctional urokinase variant M18.  
 DE  
 XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
 KW urokinase; variant; mutein.  
 XX  
 XX Synthetic.  
 OS  
 FH Key Location/Qualifiers  
 FT Region 1..365  
 FT /label= M4  
 FT /note= "unglycosylated prourokinase (Ser47-Leu411)"  
 FT Disulfide-bond 25..67  
 FT Disulfide-bond 56..80  
 FT Disulfide-bond 102..233  
 FT Disulfide-bond 143..159  
 FT Disulfide-bond 151..222  
 FT Disulfide-bond 247..316  
 FT Disulfide-bond 279..295  
 FT Disulfide-bond 306..334  
 FT Region 366..372  
 FT /label= X1  
 FT Region 373..393  
 FT /label= Y1  
 XX  
 PN DE4323754-Cl.  
 XX  
 PD 01-DEC-1994.  
 XX  
 XX 15-JUL-1993; 93DE-04323754.  
 XX  
 PR 15-JUL-1993; 93DE-04323754.  
 XX  
 XX (CHEF ) GRUENENTHAL GMBH.  
 PA  
 XX Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
 PI Saunders DJ;  
 PI



XX WPI; 1995-015191/03.  
 XX New bifunctional urokinase derivs and related plasmids - with improved  
 PT fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
 PT cerebral infarct, pulmonary embolism, etc.  
 XX  
 XX Example 1; Page 10 and Fig 1; 34pp; German.  
 XX  
 CC Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
 CC are claimed (see features table). Sequences AAR66244-R66266 are specific  
 CC examples of such derivs. which have both improved fibrinolytic and  
 CC thrombin-inhibiting activities, compared to known plasminogen activators  
 CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
 CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
 CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
 CC to correct PN field.)  
 XX  
 XX Sequence 393 AA;

Query Match 100.0%; Score 1508; DB 2; Length 393;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-125;  
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KPSSPPEELKFCQGQKTLRPFKIIGGFTTIENQPFWFAAIYRRHGGSVTVCGGSLIS 60  
 DB 90 KPSSPPEELKFCQGQKTLRPFKIIGGFTTIENQPFWFAAIYRRHGGSVTVCGGSLIS 149  
 QY 61 PCWVISATHCFIDYPKKEDYIVLGRSRLNSNTQGMKFEVENLILHKDYSADTLAHND 120  
 DB 150 PCWVISATHCFIDYPKKEDYIVLGRSRLNSNTQGMKFEVENLILHKDYSADTLAHND 209  
 QY 121 IALLKIRSKGRCAQPSRTIQTCLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT 180  
 DB 210 IALLKIRSKGRCAQPSRTIQTCLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT 269  
 QY 181 VKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240  
 DB 270 VKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 329  
 QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276  
 DB 330 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 365

RESULT 18  
 AAR66253  
 ID AAR66253 standard; protein; 393 AA.  
 XX  
 AC AAR66253;  
 XX  
 XX 25-MAR-2003 (revised)  
 DT 17-AUG-1995 (first entry)  
 XX  
 DE Bifunctional urokinase variant M20.  
 XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
 KW urokinase; variant; mutein.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..365  
 FT /label= M4  
 FT /note= "unglycosylated prourokinase (Ser47-Leu411)"  
 FT  
 FT Disulfide-bond 25..67  
 FT Disulfide-bond 56..80  
 FT Disulfide-bond 102..233  
 FT Disulfide-bond 143..159  
 FT Disulfide-bond 151..222  
 FT Disulfide-bond 247..316  
 FT Disulfide-bond 279..295

FT Disulfide-bond 306..334  
 FT Region 366..372  
 FT /label= X1  
 FT Region 373..393  
 FT /label= Y1  
 XX  
 PN DE4323754-Cl.  
 XX  
 PD 01-DEC-1994.  
 XX  
 XX 15-JUL-1993; 93DE-04323754.  
 XX  
 XX 15-JUL-1993; 93DE-04323754.  
 FR  
 XX (CHEF ) GRUENENTHAL GMBH.  
 PA  
 XX Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
 PI Saunders DJ;  
 XX  
 XX WPI; 1995-015191/03.  
 DR  
 XX New bifunctional urokinase derivs and related plasmids - with improved  
 PT fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
 PT cerebral infarct, pulmonary embolism, etc.  
 PT  
 XX Example 1; Page 10 and Fig 1; 34pp; German.

XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
 CC are claimed (see features table). Sequences AAR66244-R66266 are specific  
 CC examples of such derivs. which have both improved fibrinolytic and  
 CC thrombin-inhibiting activities, compared to known plasminogen activators  
 CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
 CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
 CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
 CC to correct PN field.)  
 XX  
 PS Sequence 393 AA;

Query Match 100.0%; Score 1508; DB 2; Length 393;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-125;  
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQGQKTLRPFKIIGGFTTIENQPFWFAAIYRRHGGSVTVCGGSLIS 60  
 DB 90 KPSSPPEELKFCQGQKTLRPFKIIGGFTTIENQPFWFAAIYRRHGGSVTVCGGSLIS 149  
 QY 61 PCWVISATHCFIDYPKKEDYIVLGRSRLNSNTQGMKFEVENLILHKDYSADTLAHND 120  
 DB 150 PCWVISATHCFIDYPKKEDYIVLGRSRLNSNTQGMKFEVENLILHKDYSADTLAHND 209  
 QY 121 IALLKIRSKGRCAQPSRTIQTCLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT 180  
 DB 210 IALLKIRSKGRCAQPSRTIQTCLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT 269  
 QY 181 VKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240  
 DB 270 VKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 329  
 QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276  
 DB 330 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 365

RESULT 19  
 AAR66249  
 ID AAR66249 standard; protein; 393 AA.  
 XX  
 AC AAR66249;  
 XX  
 XX 25-MAR-2003 (revised)  
 DT 17-AUG-1995 (first entry)  
 XX  
 DE Bifunctional urokinase variant M16.

XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
 KW urokinase; variant; mutein.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1. .365  
 FT /label= M4  
 FT /note= "unglycosylated prourokinase (Ser47-Leu411)"  
 FT Disulfide-bond 4. .85  
 FT Disulfide-bond 25. .67  
 FT Disulfide-bond 56. .80  
 FT Disulfide-bond 102. .233  
 FT Disulfide-bond 143. .159  
 FT Disulfide-bond 151. .222  
 FT Disulfide-bond 247. .316  
 FT Disulfide-bond 279. .295  
 FT Disulfide-bond 306. .334  
 FT Region /label= X1  
 FT Region 373. .393  
 FT /label= Y1  
 XX DB4323754-C1.  
 XX PN 01-DEC-1994.  
 XX PD 15-JUL-1993; 93DB-04323754.  
 XX PF 15-JUL-1993; 93DB-04323754.  
 XX PR 15-JUL-1993; 93DB-04323754.  
 XX XX (CHEF ) GRUENENTHAL GMBH.  
 XX PA Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
 PI Saunders DJ;  
 XX WPI; 1995-015191/03.  
 XX DR New bifunctional urokinase derivs and related plasmids - with improved  
 XX PT fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
 XX PT cerebral infarct, pulmonary embolism, etc.  
 XX XX Example 1; Page 10 and Fig 1; 34pp; German.  
 XX CC Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
 XX CC are claimed (see features table). Sequences AAR66244-R66266 are specific  
 XX CC examples of such derivs. which have both improved fibrinolytic and  
 XX CC thrombin-inhibiting activities, compared to known plasminogen activators  
 XX CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
 XX CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
 XX CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
 XX CC to correct PN field.)  
 XX SQ Sequence 393 AA;  
 Query Match 100.0%; Score 1508; DB 2; Length 393;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-125;  
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KPSSPPPELFCQCKTLRFRKLIIGFEFTIENQPMFAAIYRRHRGGSVYVCGSLIS 60  
 DB 90 KPSSPPPELFCQCKTLRFRKLIIGFEFTIENQPMFAAIYRRHRGGSVYVCGSLIS 149  
 QY 61 PCWVISATHCFIDYPPKEDYIVYIGRSLNSNTQGMKFEVENILHKOYSADTLAHND 120  
 DB 150 PCWVISATHCFIDYPPKEDYIVYIGRSLNSNTQGMKFEVENILHKOYSADTLAHND 209  
 QY 121 IALLKIRSKGRCAQPSRTIQTICLPWYNDPFGTSCEITGFGKNSDYLYPEQWKMT 180  
 DB 210 IALLKIRSKGRCAQPSRTIQTICLPWYNDPFGTSCEITGFGKNSDYLYPEQWKMT 269  
 QY 181 VVKLIHRECCQPHYGVSEVTTMLCAADPQWKIDSCQDSGGPLVCSLQGRMTLTGIVS 240

DB 270 VVKLIHRECCQPHYGVSEVTTMLCAADPQWKIDSCQDSGGPLVCSLQGRMTLTGIVS 329  
 QY 241 WGRGCALKKPKGYTVRVSHFLPWIRSHTEENGLAL 276  
 DB 330 WGRGCALKKPKGYTVRVSHFLPWIRSHTEENGLAL 365  
 RESULT 20  
 AAR66252  
 ID AAR66252 standard; protein; 393 AA.  
 XX AC AAR66252;  
 XX DT 25-MAR-2003 (revised)  
 DT 17-AUG-1995 (first entry)  
 XX DE Bifunctional urokinase variant M19.  
 XX KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
 KW urokinase; variant; mutein.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 FT Region 1. .365  
 FT /label= M4  
 FT /note= "unglycosylated prourokinase (Ser47-Leu411)"  
 FT Disulfide-bond 4. .85  
 FT Disulfide-bond 25. .67  
 FT Disulfide-bond 56. .80  
 FT Disulfide-bond 102. .233  
 FT Disulfide-bond 143. .159  
 FT Disulfide-bond 151. .222  
 FT Disulfide-bond 247. .316  
 FT Disulfide-bond 279. .295  
 FT Disulfide-bond 306. .334  
 FT Region /label= X1  
 FT Region 373. .393  
 FT /label= Y1  
 XX DB4323754-C1.  
 XX PN 01-DEC-1994.  
 XX PD 15-JUL-1993; 93DE-04323754.  
 XX PR 15-JUL-1993; 93DE-04323754.  
 XX XX (CHEF ) GRUENENTHAL GMBH.  
 XX PA Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
 PI Saunders DJ;  
 XX WPI; 1995-015191/03.  
 XX DR New bifunctional urokinase derivs and related plasmids - with improved  
 XX PT fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
 XX PT cerebral infarct, pulmonary embolism, etc.  
 XX XX Example 1; Page 10 and Fig 1; 34pp; German.  
 XX CC Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
 XX CC are claimed (see features table). Sequences AAR66244-R66266 are specific  
 XX CC examples of such derivs. which have both improved fibrinolytic and  
 XX CC thrombin-inhibiting activities, compared to known plasminogen activators  
 XX CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
 XX CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
 XX CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
 XX CC to correct PN field.)  
 XX SQ Sequence 393 AA;

DR WPI; 1995-015191/03.

XX New bifunctional urokinase derivs and related plasmids - with improved

PT fibrinolytic and thrombin inhibiting activities, for treating cardiac and

PT cerebral infarct, pulmonary embolism, etc.

XX Example 1; Page 10 and Fig 1; 34pp; German.

PS Bifunctional urokinase derivatives corresponding to the formula M4-X1-V1

XX are claimed (see features table). Sequences AAR6244-R66266 are specific

CC examples of such derivs. which have both improved fibrinolytic and

CC thrombin-inhibiting activities, compared to known plasminogen activators

CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,

CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac

CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003

XX to correct PN field.)

XX Sequence 393 AA;

Query Match 100.0%; Score 1508; DB 2; Length 393;

Best Local Similarity 100.0%; Pred. No. 2.7e-125; Indels 0; Gaps 0

Matches 276; Conservative 0; Mismatches 0;

Qy 1 KPSSPPEELKFCQGQKTLRPRFKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGSLIS 60

Db 90 KPSSPPEELKFCQGQKTLRPRFKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGSLIS 149

Qy 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKEFEVENLILHKDYSADTLAHND 120

Db 150 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKEFEVENLILHKDYSADTLAHND 209

Qy 121 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMT 190

Db 210 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMT 269

Qy 181 VKLISHRECQPHYVGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRWTLTGIVS 240

Db 270 VKLISHRECQPHYVGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRWTLTGIVS 329

Qy 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276

Db 330 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 365

RESULT 22

AAR66250

ID AAR66250 standard; protein; 393 AA.

XX AAR66250;

XX 25-MAR-2003 (revised)

DT 17-AUG-1995 (first entry)

XX Bifunctional urokinase variant M17.

XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;

KW urokinase; variant; mutein.

XX Synthetic.

XX Key Location/Qualifiers

PH Region 1. .365

FT /label= M4

FT /note= "unglycosylated prourokinase (Ser47-Leu411)"

FT Disulfide-bond 4. .85

FT Disulfide-bond 25. .67

FT Disulfide-bond 56. .80

FT Disulfide-bond 102. .233

FT Disulfide-bond 143. .159

FT Disulfide-bond 151. .222

FT Disulfide-bond 247. .316

FT Disulfide-bond 279. .295

FT Disulfide-bond 306. .334

Query Match 100.0%; Score 1508; DB 2; Length 393;

Best Local Similarity 100.0%; Pred. No. 2.7e-125; Indels 0; Gaps 0;

Matches 276; Conservative 0; Mismatches 0;

Qy 1 KPSSPPEELKFCQGQKTLRPRFKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGSLIS 60

Db 90 KPSSPPEELKFCQGQKTLRPRFKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGSLIS 149

Qy 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKEFEVENLILHKDYSADTLAHND 120

Db 150 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKEFEVENLILHKDYSADTLAHND 209

Qy 121 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMT 190

Db 210 IALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMT 269

Qy 181 VKLISHRECQPHYVGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRWTLTGIVS 240

Db 270 VKLISHRECQPHYVGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRWTLTGIVS 329

Qy 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276

Db 330 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 365

RESULT 21

AAR66244

ID AAR66244 standard; protein; 393 AA.

XX AAR66244;

XX 25-MAR-2003 (revised)

DT 22-AUG-1995 (first entry)

XX Bifunctional urokinase variant M11.

XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;

KW urokinase; variant; mutein.

XX Synthetic.

XX Key Location/Qualifiers

PH Region 1. .365

FT /label= M4

FT /note= "unglycosylated prourokinase (Ser47-Leu411)"

FT Disulfide-bond 4. .85

FT Disulfide-bond 25. .67

FT Disulfide-bond 56. .80

FT Disulfide-bond 102. .233

FT Disulfide-bond 143. .159

FT Disulfide-bond 151. .222

FT Disulfide-bond 247. .316

FT Disulfide-bond 279. .295

FT Disulfide-bond 306. .334

FT Region 366. .374

FT /label= X1

FT Region 375. .393

FT /label= Y1

XX DE4323754-C1.

XX 01-DEC-1994.

XX 15-JUL-1993; 93DE-04323754.

XX 15-JUL-1993; 93DE-04323754.

XX (CHEF ) GRUENENTHAL GMBH.

XX Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;

PI Saunders DJ;

XX

plasma

angina

Synthetic.

EP714982-A2.

05-JUN-1996.

16-NOV-1995; 95BP-00118050.

30-NOV-1994; 94DE-04442665.

(CHEF ) GRUENTHAL GMBH.

Wendt S, Steffens GJ, Janocha E, Heinzel-Wieland R;  
WPI; 1996-269715/28.

Chimeric protein contg. plasminogen activating sequence and thrombin-inhibiting sequence - useful as thrombus-specific thrombolytic agent with rapid action.

Example 1; Page 21-22; 37pp; German.

Example 1 describes the prodn. of plasmids pSE1 and pSE9 contg. a DNA encoding a chimeric protein with fibrinolytic and thrombin-inhibiting properties. pSE1 encodes the protein given in AAR99597 and pSE9 encodes the protein given in AAR99596

Sequence 393 AA;

Query Match 100.0%; Score 1508; DB 2; Length 393;  
Best Local Similarity 100.0%; Pred. No. 2.7e-125;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 KPSSPPEELKFCQGCKTLPRFKIIGGEFTTIENQPWFPAAIYRRHRGGSVTVVCGGSLIS 60  
DB 91 KPSSPPEELKFCQGCKTLPRFKIIGGEFTTIENQPWFPAAIYRRHRGGSVTVVCGGSLIS 150  
QY 61 PCWVISATHCFFIDYPKKEDYIVVLGRSRLNSNTQGMKFVEVENLILHKDYSADTLLAHND 120  
DB 151 PCWVISATHCFFIDYPKKEDYIVVLGRSRLNSNTQGMKFVEVENLILHKDYSADTLLAHND 210  
QY 121 IALLKIRSGRCAQPSRTIQITCLPSMYNDPQFGTSCEITGFGENSTDYLYPEQLKMT 180  
DB 211 IALLKIRSGRCAQPSRTIQITCLPSMYNDPQFGTSCEITGFGENSTDYLYPEQLKMT 270  
QY 181 VKKLISHRECCQPHYGVSEVTVMKLCADPOWKTDSCQDSGGPLVCISLQGRMTLTGIS 240  
DB 271 VKKLISHRECCQPHYGVSEVTVMKLCADPOWKTDSCQDSGGPLVCISLQGRMTLTGIS 330  
QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTEENGLAL 276  
DB 331 WGRGCALKDKPGVYTRVSHFLPWIRSHTEENGLAL 366

RESULT 24  
AAR99596  
ID AAR99597 standard; protein; 393 AA.  
XX AC AAR99596;  
DT DT 05-DEC-1996 (first entry)  
XX XX Chimeric protein M37 encoded by pSE9.  
XX KW Thrombin; inhibition; thrombus; thrombolysis; chimeric protein;  
KW plasma activating sequence; fibrinolysis; infarction;  
XX angina pectoris; deep vein thrombosis.  
XX Synthetic.



XX Bifunctional urokinase variant M32.  
 DE fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
 KW urokinase; variant; mutein.  
 XX Synthetic.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1. .365  
 FT /label= M4  
 FT /note= "unglycosylated prourokinase (Ser47-Leu411)"  
 FT Disulfide-bond 25. .67  
 FT Disulfide-bond 56. .80  
 FT Disulfide-bond 102. .233  
 FT Disulfide-bond 143. .159  
 FT Disulfide-bond 151. .222  
 FT Disulfide-bond 247. .316  
 FT Disulfide-bond 279. .295  
 FT Disulfide-bond 306. .334  
 FT Region 366. .371  
 FT /label= X1  
 FT Region 372. .395  
 FT /label= Y1  
 XX DE4323754-Cl.  
 XX  
 PD 01-DEC-1994.  
 XX  
 PF 15-JUL-1993; 93DE-04323754.  
 XX  
 PR 15-JUL-1993; 93DE-04323754.  
 XX  
 XX (CHEF ) GRUENENTHAL GMBH.  
 XX  
 XX Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
 PI Saunders DJ;  
 XX  
 XX WPI; 1995-015191/03.  
 XX  
 XX New bifunctional urokinase derivs and related plasmids - with improved  
 PT fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
 PT cerebral infarct, pulmonary embolism, etc.  
 XX  
 XX Example 1; Page 11 and Fig 1; 34pp; German.  
 XX  
 XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
 CC are claimed (see features table). Sequences AAR66244-R66266 are specific  
 CC examples of such derivs. which have both improved fibrinolytic and  
 CC thrombin-inhibiting activities, compared to known plasminogen activators  
 CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
 CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
 CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
 CC to correct PN field.)  
 XX  
 SQ Sequence 395 AA;  
 Query Match 100.0%; Score 1508; DB 2; Length 395;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-125;  
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KPSSPEELKFCQGKTLPRFKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGSLIS 60  
 DB 90 KPSSPEELKFCQGKTLPRFKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGSLIS 149  
 QY 61 PCWVISATCFIDYPKEDYIVVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHDND 120  
 DB 150 PCWVISATCFIDYPKEDYIVVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHDND 209  
 QY 121 IALLKRSKRGCAQPSRTITQICLPSMYNDPQFGTSCEITGFGKENSIDYILPQLKMT 180  
 DB 210 IALLKRSKRGCAQPSRTITQICLPSMYNDPQFGTSCEITGFGKENSIDYILPQLKMT 269

QY 181 VVKLISHRECQPHYVGSEVTTKMLCAADPOWKTSCQDSCGPLVCSLQGRMTLTGIVS 240  
 DB 270 VVKLISHRECQPHYVGSEVTTKMLCAADPOWKTSCQDSCGPLVCSLQGRMTLTGIVS 329  
 QY 241 WGRGKALDKPGVYTRVSHFLPWIRSHTKENGLAL 276  
 DB 330 WGRGKALDKPGVYTRVSHFLPWIRSHTKENGLAL 365  
 RESULT 27  
 AAR66262  
 ID AAR66262 standard; protein; 395 AA.  
 XX  
 AC AAR66262;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 17-AUG-1995 (first entry)  
 XX  
 DE Bifunctional urokinase variant M29.  
 XX  
 XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;  
 KW urokinase; variant; mutein.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1. .365  
 FT /label= M4  
 FT /note= "unglycosylated prourokinase (Ser47-Leu411)"  
 FT Disulfide-bond 25. .67  
 FT Disulfide-bond 56. .80  
 FT Disulfide-bond 102. .233  
 FT Disulfide-bond 143. .159  
 FT Disulfide-bond 151. .222  
 FT Disulfide-bond 247. .316  
 FT Disulfide-bond 279. .295  
 FT Disulfide-bond 306. .334  
 FT Region 366. .371  
 FT /label= X1  
 FT Region 372. .395  
 FT /label= Y1  
 XX DE4323754-Cl.  
 XX  
 PD 01-DEC-1994.  
 XX  
 PF 15-JUL-1993; 93DE-04323754.  
 XX  
 PR 15-JUL-1993; 93DE-04323754.  
 XX  
 XX (CHEF ) GRUENENTHAL GMBH.  
 XX  
 XX Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;  
 PI Saunders DJ;  
 XX  
 XX WPI; 1995-015191/03.  
 XX  
 XX New bifunctional urokinase derivs and related plasmids - with improved  
 PT fibrinolytic and thrombin inhibiting activities, for treating cardiac and  
 PT cerebral infarct, pulmonary embolism, etc.  
 XX  
 XX Example 1; Page 11 and Fig 1; 34pp; German.  
 XX  
 XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1  
 CC are claimed (see features table). Sequences AAR66244-R66266 are specific  
 CC examples of such derivs. which have both improved fibrinolytic and  
 CC thrombin-inhibiting activities, compared to known plasminogen activators  
 CC or thrombin inhibitors. The proteins are useful as thrombolytic agents,  
 CC e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac  
 CC and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003  
 CC to correct PN field.)

XX	Sequence 395 AA;	
XX	Query Match	100.0%; Score 1508; DB 2; Length 395;
XX	Best Local Similarity	100.0%; Pred. No. 2.7e-125;
XX	Matches 276; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 KPSSPEELKFOGQKTLRPRFKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGSLIS	60
DB	90 KPSSPEELKFOGQKTLRPRFKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGSLIS	149
QY	61 PCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND	120
DB	150 PCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND	209
QY	121 IALLKIRSEKGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT	180
DB	210 IALLKIRSEKGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT	269
QY	181 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS	240
DB	270 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS	329
QY	241 WGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL	276
DB	330 WGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL	365
RESULT 28		
AAR66246		
ID	AAR66246 standard; protein; 396 AA.	
AC	AAR66246;	
XX		
XX	25-MAR-2003 (revised)	
DT	17-AUG-1995 (first entry)	
XX		
XX	Bifunctional urokinase variant M13.	
DE		
XX	fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;	
KW	urokinase; variant; mutein.	
XX		
OS	Synthetic.	
XX		
XX	Key	Location/Qualifiers
FT	Region	1..365
FT		/label= M4
FT		/note= "unglycosylated prourokinase (Ser47-Leu411)"
FT	Disulfide-bond	4..85
FT	Disulfide-bond	25..67
FT	Disulfide-bond	56..80
FT	Disulfide-bond	102..233
FT	Disulfide-bond	143..159
FT	Disulfide-bond	247..316
FT	Disulfide-bond	279..295
FT	Disulfide-bond	306..334
FT	Region	366..377
FT		/label= X1
FT	Region	378..396
FT		/label= Y1
XX		
PN	DE4323754-C1.	
XX		
PD	01-DEC-1994.	
XX		
PF	15-JUL-1993; 93DE-04323754.	
XX		
PR	15-JUL-1993; 93DE-04323754.	
XX		
PA	(CHEF ) GRUENENTHAL GMBH.	
XX		
PI	Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;	

PI	Saunders DJ;	
XX		
DR	WPI; 1995-015191/03.	
XX		
PT	New bifunctional urokinase derivs and related plasmids - with improved	
PT	fibrinolytic and thrombin inhibiting activities, for treating cardiac and	
PT	cerebral infarct, pulmonary embolism, etc.	
XX		
PS	Example 1; Page 10 and Fig 1; 34pp; German.	
XX		
CC	Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1	
CC	are claimed (see features table). Sequences AAR66244-R66266 are specific	
CC	examples of such derivs. which have both improved fibrinolytic and	
CC	thrombin-inhibiting activities, compared to known plasminogen activators	
CC	or thrombin inhibitors. The proteins are useful as thrombolytic agents,	
CC	e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac	
CC	and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003	
CC	to correct FN field.)	
XX		
SQ	Sequence 396 AA;	
	Query Match	100.0%; Score 1508; DB 2; Length 396;
	Best Local Similarity	100.0%; Pred. No. 2.7e-125;
	Matches 276; Conservative	0; Mismatches 0; Indels 0; Gaps 0.
QY	1 KPSSPEELKFOGQKTLRPRFKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGSLIS	60
DB	90 KPSSPEELKFOGQKTLRPRFKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGSLIS	149
QY	61 PCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND	120
DB	150 PCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND	209
QY	121 IALLKIRSEKGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT	180
DB	210 IALLKIRSEKGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT	269
QY	181 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS	240
DB	270 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS	329
QY	241 WGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL	276
DB	330 WGRGCAKDKPGVYTRVSHFLPWIRSHTKENGLAL	365
RESULT 29		
AAR66248		
ID	AAR66248 standard; protein; 397 AA.	
XX		
AC	AAR66248;	
XX		
DT	25-MAR-2003 (revised)	
DT	17-AUG-1995 (first entry)	
XX		
DE	Bifunctional urokinase variant M15.	
XX		
KW	fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;	
KW	urokinase; variant; mutein.	
XX		
OS	Synthetic.	
XX		
XX	Key	Location/Qualifiers
FT	Region	1..365
FT		/label= M4
FT		/note= "unglycosylated prourokinase (Ser47-Leu411)"
FT	Disulfide-bond	4..85
FT	Disulfide-bond	25..67
FT	Disulfide-bond	56..80
FT	Disulfide-bond	102..233
FT	Disulfide-bond	143..159
FT	Disulfide-bond	247..316
FT	Disulfide-bond	279..295
FT	Disulfide-bond	306..334
FT	Region	366..377
FT		/label= X1
FT	Region	378..396
FT		/label= Y1
XX		
PN	DE4323754-C1.	
XX		
PD	01-DEC-1994.	
XX		
PF	15-JUL-1993; 93DE-04323754.	
XX		
PR	15-JUL-1993; 93DE-04323754.	
XX		
PA	(CHEF ) GRUENENTHAL GMBH.	
XX		
PI	Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;	

```

FT Disulfide-bond 279. .295
FT Disulfide-bond 306. .334
FT Region 366. .378
FT /label= X1
FT Region 379. .397
FT /label= Y1
XX
PN DE4323754-Cl.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-04323754.
XX
XX 15-JUL-1993; 93DE-04323754.
XX
XX (CHEF ) GRUENTHAL GMBH.
XX
XX Steffens GJ, Wnendt S, Schneider J, Heinzel-Wieland R;
XX Saunders DJ;
XX WFI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with improved
XX fibrinolytic and thrombin inhibiting activities, for treating cardiac and
XX cerebral infarct, pulmonary embolism, etc.
XX
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1
XX are claimed (see features table). Sequences AAR66244-R66266 are specific
XX examples of such derivs, which have both improved fibrinolytic and
XX thrombin-inhibiting activities, compared to known plasminogen activators
XX or thrombin inhibitors. The proteins are useful as thrombolytic agents,
XX e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac
XX and cerebral infarction and pulmonary embolism. (Updated on 25-MAR-2003
XX to correct PN field.)
XX
XX Sequence 397 AA;
XX
XX Query Match 100.0%; Score 1508; DB 2; Length 397;
XX Best Local Similarity 100.0%; Pred. No. 2.7e-125;
XX Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0
XX
Qy 1 KPSPPEELKFCGGQKTLRPRFKIIGGFETTINQNPWFAAIYRRHGGSVYVCGSLIS 60
Db 90 KPSPPEELKFCGGQKTLRPRFKIIGGFETTINQNPWFAAIYRRHGGSVYVCGSLIS 149
Qy 61 PCWVISATHCFDYPKEDYIVYLGSRSLNNTGEMKFEVENILHKKOYSADTLAHND 120
Db 150 PCWVISATHCFDYPKEDYIVYLGSRSLNNTGEMKFEVENILHKKOYSADTLAHND 209
Qy 121 IALLKIRSGRCAQPSRTIQICLPSPMYNDPQGTSCETGFGKENSTDYLYPEQLKMT 180
Db 210 IALLKIRSGRCAQPSRTIQICLPSPMYNDPQGTSCETGFGKENSTDYLYPEQLKMT 269
Qy 181 VWKLISHRECCQPHYGSSVTTTMLCAADPWKTDSCQSGGGLVCSLOGRMTLTGIVS 240
Db 270 VWKLISHRECCQPHYGSSVTTTMLCAADPWKTDSCQSGGGLVCSLOGRMTLTGIVS 329
Qy 241 WGRGCALKDPGYVTVRVSHFLPWIRSHTKENGLAL 276
Db 330 WGRGCALKDPGYVTVRVSHFLPWIRSHTKENGLAL 365
XX
XX RESULT 30
XX AAP50871
XX ID AAP50871 standard; protein; 411 AA.
XX AC AAP50871;
XX
XX DT 30-NOV-1991 (first entry)
XX
XX Sequence encoded by cDNA sequence for human urokinase zymogen (Japanese

```



DB	376	WGRGCAKDKPGVYTRVSHFLPWIRSHTKEENGLAL	411
RESULT 31			
ID	AAP91684	AAP91684 standard; protein; 411 AA.	
AC	AAP91684;		
DT	25-MAR-2003 (revised)		
DT	13-MAR-1992 (first entry)		
DE	Sequence of urokinase.		
KW	Plasminogen activator; urokinase; thrombosis therapy; cardiovascular disorder; myocardial infarct.		
OS	Homo sapiens.		
PH	Key	Location/Qualifiers	
FT	Disulfide-bond	13..31	
FT	Disulfide-bond	19..11	
FT	Disulfide-bond	33..42	
FT	Disulfide-bond	50..131	
FT	Disulfide-bond	71..113	
FT	Disulfide-bond	102..126	
FT	Disulfide-bond	148..279	
FT	Disulfide-bond	189..205	
FT	Disulfide-bond	197..268	
FT	Disulfide-bond	293..361	
FT	Modified-site	302	
FT	/label= glycosylation site		
FT	Disulfide-bond	325..341	
FT	Disulfide-bond	352..380	
XX	WO8911531-A.		
XX	30-NOV-1989.		
XX	04-MAY-1989;	89WO-US001947.	
XX	20-MAY-1988;	88US-00196909.	
XX	(GETH ) GENENTECH INC.		
XX	Anderson S, Keyt B;		
XX	WPI; 1989-370725/50.		
PT	New plasminogen activator variants with additional glycosylation - having increased circulating half life in plasma, used for treating cardiovascular disorders.		
PS	Disclosure; Fig 3; 60pp; English.		
CC	The new plasminogen activator (PA) variants of the invention are derived from human urokinase, prourokinase or esp. tissue PA (tPA). They are useful for treating cardiovascular disorders such as myocardial infarct or thrombosis. Typical doses are 0.3 mg/kg for infarct and 0.1-0.2 mg/kg for thrombosis. (Updated on 25-MAR-2003 to correct PF field.)		
XX	Sequence 411 AA;		
SQ	Query Match 100.0%; Score 1508; DB 1; Length 411; Best Local Similarity 100.0%; Pred. No. 2.8e-125; Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	KPSSPPBELKFCQCKTLRPRFKIIGGEFTTIENQNPWFAAIYRRHRGGSVTVYCGGSLIS	60
DB	136	KPSSPPBELKFCQCKTLRPRFKIIGGEFTTIENQNPWFAAIYRRHRGGSVTVYCGGSLIS	195
QY	61	PCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENILHKDYSADTLAHND	120
DB	196	PCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENILHKDYSADTLAHND	255
QY	121	IALLKIRSEKRCQAPSRITQIICLPMSYNDPQGTSCITGFGKENSTDYLYPEQLKMT	180
DB	256	IALLKIRSEKRCQAPSRITQIICLPMSYNDPQGTSCITGFGKENSTDYLYPEQLKMT	315

QY 181 VVKLISHRECOQPHYHGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240  
 DB 316 VVKLISHRECOQPHYHGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 375  
 QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 276  
 DB 376 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411

## RESULT 33

AAP93589  
 ID AAP93589 standard; protein; 411 AA.

AC AAP93589;  
 XX

DT 19-JUN-1990 (first entry)

XX Amino acid sequence of single-chain pro-urokinase extracted from human renal cells.

DE Single-chain pro-urokinase; fibrinolytic activity enhancer; plasminogen; thrombosis.

OS Homo sapiens.

PN EP310065-A.

PD 05-APR-1989.

XX 29-SEP-1989; 88EP-00116067.

PR 01-OCT-1987; 87JP-00248937.

XX (GREG ) GREEN CROSS CORP.

PI Teukada M, Tanaka K, Iga Y;

XX WPI; 1989-101389/14.

XX Fibrinolytic activity enhancer - comprising plasminogen for enhancing activity of single-chain pro-urokinase without causing systemic fibrinolysis.

PS Disclosure; Page 7; 8pp; English.

XX It was extracted from human renal cells and purified using a monoclonal antibody for single-chain pro-urokinase (scpu) recovered from hybridoma cells created from the fusion of mouse myeloma cells with mouse BALB/c spleen cells which had been preliminarily immunised with scpu. The purified product has a molecular weight of 54,000 in SDS/polyacrylamide gel electrophoresis. The combined use of scpu and plasminogen at a specific ratio can enhance fibrinolytic activity of scpu without causing any systemic fibrinolysis. This enables lowering the dose of scpu which might relieve side effects. They can be used for the treatment of thrombosis and obstructive diseases

XX Sequence 411 AA;

Query Match 100.0%; Score 1508; DB 1; Length 411;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-125;  
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSPPEELKFCQCKTLRPRFKIIGGEFTTIENQPFPAIYRRHGGSVTVVCGSLIS 60  
 DB 136 KPSPPEELKFCQCKTLRPRFKIIGGEFTTIENQPFPAIYRRHGGSVTVVCGSLIS 195  
 QY 61 PCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120  
 DB 196 PCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 255  
 QY 121 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMT 180

DB 256 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMT 315  
 QY 181 VVKLISHRECOQPHYHGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240  
 DB 316 VVKLISHRECOQPHYHGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 375  
 QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 276  
 DB 376 WGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411

## RESULT 34

AAR06244  
 ID AAR06244 standard; protein; 411 AA.

XX AAR06244;  
 AC

DT 07-DEC-1990 (first entry)

XX Urokinase precursor protein.

XX Urokinase precursor; fibrinolysis; thrombolytic; cerebral thrombosis; myocardial infarction.

OS Homo sapiens.

PN EP380334-A.

PD 01-AUG-1990.

XX 25-JAN-1990; 90EP-00300772.

PR 27-JAN-1989; 89JP-00016406.

XX 17-MAY-1989; 89JP-00121405.

XX (GREG ) GREEN CROSS CORP.

PI Matsuda H, Ueda Y, Tamanouchi K;

XX WPI; 1990-233117/31.

XX Urokinase precursor-lipid composite - used as thrombolytic agent, having prolonged half-life in the blood, enhanced bio-availability and improved activity.

PS Claim 3; Fig 1; 11pp; English.

XX By forming a precursor-lipid composite, the half-life of this thrombolytic agent in the blood may be increased, exhibiting improved activity without abnormal acceleration of fibrinolytic activity. Compound is useful as a thrombolytic agent in treatment of cerebral thrombosis, myocardial infarction etc

XX Sequence 411 AA;

Query Match 100.0%; Score 1508; DB 2; Length 411;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-125;  
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 KPSPPEELKFCQCKTLRPRFKIIGGEFTTIENQPFPAIYRRHGGSVTVVCGSLIS 60  
 DB 136 KPSPPEELKFCQCKTLRPRFKIIGGEFTTIENQPFPAIYRRHGGSVTVVCGSLIS 195  
 QY 61 PCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120  
 DB 196 PCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 255  
 QY 121 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMT 180  
 DB 256 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMT 315  
 QY 181 VVKLISHRECOQPHYHGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240

```
Db 316 VVKLISHRECCQPHYYGSEVTTKMLCAADPWKTDSCQDGGPLVCSLQGRMTLTGIVS 375
QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
Db 376 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411

RESULT 35
AAR07904
XX AAR07904 standard; protein; 411 AA.
AC AAR07904;
XX 21-FEB-1991 (first entry)
XX Human pro-urokinase variant.
DE Thrombin; fibrin; bleeding; pHR27.
XX Homo sapiens.
FH Key Location/Qualifiers
FT Domain 10..42
FT /label= Epidermal growth factor (EGF) domain
FT Region 10..19
FT /label= First loop
FT Region 20..31
FT /label= Second loop
FT Active-site 27..29
FT /label= Modified site
FT Region 33..42
FT /label= Third loop
XX EP398362-A.
XX 22-NOV-1990.
XX 18-MAY-1990; 90EP-00109473.
XX 18-MAY-1989; 89JP-00126434.
XX (GREC ) GREEN CROSS CORP.
XX Tanabe T, Amatsuji Y, Kasai S, Hirose M, Morita M, Kawabe H;
XX Arimura H;
XX WPI; 1990-350147/47.
XX N-PSDB; AAQ06135.
XX Human pro-urokinase variant - produced by recombinant methods, showing
XX increased half life in blood and high affinity for fibrin.
XX Disclosure; Fig 1; 27pp; English.
XX Modified pro-urokinase has a longer half-life in blood, and dissolves
XX thrombin without causing the spontaneous bleeding associated with
XX urokinase. The modification puts an epidermal growth factor domain into
XX the protein, containing the sequence Asn-X-Ser or Asn-X-Thr where X is
XX any residue. Plasmid pHR27 is disclosed as containing the modified
XX sequence
XX Sequence 411 AA;
Query Match 100.0%; Score 1508; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.8e-125; Indels 0; Gaps 0;
Matches 276; Conservative 0; Mismatches 0;

QY 1 KPSSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQPFWFAIYRRHGGSVTVVCGSLIS 60
Db 136 KPSSPPEELKFCQCGKTLRPRFKIIGGEFTTIENQPFWFAIYRRHGGSVTVVCGSLIS 195
QY 61 PCWVISATHCFIDYPKKEDYIVVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
```

```
Db 196 PCWVISATHCFIDYPKKEDYIVVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 255
QY 121 IALLKIRSKEGRCQAQPSRTIQTICLPSMYNDPOFGTSCBITGPKENSTDYLYPEQLKWT 180
Db 256 IALLKIRSKEGRCQAQPSRTIQTICLPSMYNDPOFGTSCBITGPKENSTDYLYPEQLKWT 315
QY 181 VVKLISHRECCQPHYYGSEVTTKMLCAADPWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
Db 316 VVKLISHRECCQPHYYGSEVTTKMLCAADPWKTDSCQDGGPLVCSLQGRMTLTGIVS 375
QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
Db 376 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411

RESULT 36
AAR07902
ID AAR07902 standard; protein; 411 AA.
XX AAR07902;
XX 21-FEB-1991 (first entry)
XX Human pro-urokinase variant.
DE Thrombin; fibrin; bleeding; pHR22.
XX Homo sapiens.
FH Key Location/Qualifiers
FT Domain 10..42
FT /label= Epidermal growth factor (EGF) domain
FT Region 10..19
FT /label= First loop
FT Region 20..31
FT /label= Second loop
FT Active-site 22..24
FT /label= Modified site
FT Region 33..42
FT /label= Third loop
XX EP398362-A.
XX 22-NOV-1990.
XX 18-MAY-1990; 90EP-00109473.
XX 18-MAY-1989; 89JP-00126434.
XX (GREC ) GREEN CROSS CORP.
XX Tanabe T, Amatsuji Y, Kasai S, Hirose M, Morita M, Kawabe H;
XX Arimura H;
XX WPI; 1990-350147/47.
XX N-PSDB; AAQ06133.
XX Human pro-urokinase variant - produced by recombinant methods, showing
XX increased half life in blood and high affinity for fibrin.
XX Disclosure; Fig 1; 27pp; English.
XX Modified pro-urokinase has a longer half-life in blood, and dissolves
XX thrombin without causing the spontaneous bleeding associated with
XX urokinase. The modification puts an epidermal growth factor domain into
XX the protein, containing the sequence Asn-X-Ser or Asn-X-Thr where X is
XX any residue. Plasmid pHR22 is disclosed as containing the modified
XX sequence
XX Sequence 411 AA;
Query Match 100.0%; Score 1508; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.8e-125;
```

Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 KPSSPPEELKFCGQKTLRPRFKIIGGEFTTIENQFWFAAIYRRHRGGSVTVVCGSLIS 60
Db	136 KPSSPPEELKFCGQKTLRPRFKIIGGEFTTIENQFWFAAIYRRHRGGSVTVVCGSLIS 195
Qy	61 PCWVISATHCFIDYPKKEDIYVILGRSRLNSNTQGMKFEVENLILHKDYSADTLAHND 120
Db	196 PCWVISATHCFIDYPKKEDIYVILGRSRLNSNTQGMKFEVENLILHKDYSADTLAHND 255
Qy	121 IALLKIRSEGRCAQPSRTIOTICLPSMYNDPQFGTSCBITGFGKENSTDYLYPEQLKMT 180
Db	256 IALLKIRSEGRCAQPSRTIOTICLPSMYNDPQFGTSCBITGFGKENSTDYLYPEQLKMT 315
Qy	181 VVKLISHRECCQPHYVGSVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
Db	316 VVKLISHRECCQPHYVGSVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 375
Qy	241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
Db	376 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411
RESULT 37	
AAR07903	
ID	AAR07903 standard; protein; 411 AA.
XX	AC AAR07903;
XX	DT 21-FEB-1991 (first entry)
DE	Human pro-urokinase variant.
XX	XX Thrombin; fibrin; bleeding; pHR24.
XX	OS Homo sapiens.
PH	Key Location/Qualifiers
FT	Domain 10..42
FT	/label= Epidermal growth factor (EGF) domain
FT	Region 10..19
FT	/label= First loop
FT	Region 20..31
FT	/label= Second loop
FT	Active-site 24..26
FT	/label= Modified site
FT	Region 33..42
FT	/label= Third loop
XX	EP398362-A.
XX	22-NOV-1990.
XX	18-MAY-1990; 90BF-00109473.
XX	18-MAY-1989; 89JP-00126434.
XX	(GREC ) GREEN CROSS CORP.
XX	Tanabe T, Amatsuji Y, Kasai S, Hirose M, Morita M, Kawabe H;
XX	Arimura H;
XX	WPI; 1990-350147/47.
XX	N-PSDB; AAQ06134.
XX	Human pro-urokinase variant - produced by recombinant methods, showing
XX	increased half life in blood and high affinity for fibrin.
XX	Disclosure; Fig 1; 27pp; English.
XX	Modified pro-urokinase has a longer half-life in blood, and dissolves
XX	thrombin without causing the spontaneous bleeding associated with
XX	urokinase. The modification puts an epidermal growth factor domain into

CC the protein, containing the sequence Asn-X-Ser or Asn-X-Thr where X is	
any residue. Plasmid pHR24 is disclosed as containing the modified	
sequence	
XX	Sequence 411 AA;
Query Match 100.0%; Score 1508; DB 2; Length 411;	
Best Local Similarity 100.0%; Pred. No. 2.8e-125;	
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
Qy	1 KPSSPPEELKFCGQKTLRPRFKIIGGEFTTIENQFWFAAIYRRHRGGSVTVVCGSLIS 60
Db	136 KPSSPPEELKFCGQKTLRPRFKIIGGEFTTIENQFWFAAIYRRHRGGSVTVVCGSLIS 195
Qy	61 PCWVISATHCFIDYPKKEDIYVILGRSRLNSNTQGMKFEVENLILHKDYSADTLAHND 120
Db	196 PCWVISATHCFIDYPKKEDIYVILGRSRLNSNTQGMKFEVENLILHKDYSADTLAHND 255
Qy	121 IALLKIRSEGRCAQPSRTIOTICLPSMYNDPQFGTSCBITGFGKENSTDYLYPEQLKMT 180
Db	256 IALLKIRSEGRCAQPSRTIOTICLPSMYNDPQFGTSCBITGFGKENSTDYLYPEQLKMT 315
Qy	181 VVKLISHRECCQPHYVGSVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
Db	316 VVKLISHRECCQPHYVGSVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 375
Qy	241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
Db	376 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411
RESULT 38	
AAR34584	
ID	AAR34584 standard; protein; 411 AA.
XX	AC AAR34584;
XX	DT 25-MAR-2003 (revised)
XX	DT 14-SEP-1993 (first entry)
XX	DE Mutant human prourokinase.
XX	KW pUK; increased half life; improved fibrin affinity.
XX	OS Homo sapiens.
XX	PN EP541952-A1.
XX	PD 19-MAY-1993.
XX	PP 06-OCT-1992; 92EP-00117000.
XX	PR 07-OCT-1991; 91JP-00289257.
XX	FA (GREC ) GREEN CROSS CORP.
XX	PI Tanabe T, Morita M, Hirose M, Amatsuji Y;
XX	DR WPI; 1993-160551/20.
XX	DR N-PSDB; AAQ41450.
XX	New human pro-urokinase mutants with thrombolytic activity - have a
XX	neutral aminoacid in the epidermal growth factor region replaced with a
XX	basic aminoacid, or an acid residue replaced with a non-acidic residue.
XX	Claim 1; Page 17-20; 38pp; English.
XX	The sequence is that of a mutant human prourokinase, in which a neutral
XX	amino acid in the epidermal growth region has been replaced with a basic
XX	amino acid, or an acidic amino acid has been replaced by a non-acidic
XX	amino acid. Preferred replacements are 16Gly->Lys, 38Gly->Lys and 45Asp-
XX	>Asn. The mutant has an increased half-life in blood as compared to the
XX	prior art mutant with a deleted EGF region. It has improved affinity for

CC fibrin, and has other features the same as human prourokinase. (Updated  
XX on 25-MAR-2003 to correct PN field.)  
SQ Sequence 411 AA;

Query Match 100.0%; Score 1508; DB 2; Length 411;  
Best Local Similarity 100.0%; Pred. No. 2.8e-125;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KPSPPPEELKFCGQKTLPRPKLIIGGEFTTIENQFWFAAIYRRHGGSVTVVCGSLIS 60  
DB 136 KPSPPPEELKFCGQKTLPRPKLIIGGEFTTIENQFWFAAIYRRHGGSVTVVCGSLIS 195  
QY 61 PCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120  
DB 196 PCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 255  
QY 121 IALLKIRSKGRCAQPSRTIOTICLPMSYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT 180  
DB 256 IALLKIRSKGRCAQPSRTIOTICLPMSYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT 315  
QY 181 VVKLISHRECOQPHYVGSVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240  
DB 316 VVKLISHRECOQPHYVGSVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 375  
QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276  
DB 376 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411

RESULT 39  
AAR62991  
ID AAR62991 standard; protein; 411 AA.

AC AAR62991;  
DT 25-MAR-2003 (revised)  
DT 21-SEP-1995 (first entry)  
XX Pro-urokinase.  
DE Pro-urokinase; thrombolysis; fibrin clot lysis.  
KW Pro-urokinase; thrombolysis; fibrin clot lysis.  
XX Homo sapiens.

XX Key Location/Qualifiers  
FT Disulfide-bond 11..19  
FT Disulfide-bond 13..31  
FT Disulfide-bond 33..42  
FT Disulfide-bond 50..131  
FT Disulfide-bond 71..113  
FT Disulfide-bond 102..126  
FT Disulfide-bond 148..279  
FT Disulfide-bond 189..205  
FT Disulfide-bond 197..268  
FT Disulfide-bond 293..362  
FT Domain 297..313  
FT /note= "flexible loop"  
FT Disulfide-bond 325..341  
FT Disulfide-bond 352..380  
XX WO9501427-A1.  
XX 12-JAN-1995.  
XX 28-JUN-1994; 94WO-US007278.  
XX 02-JUL-1993; 93US-00087163.  
XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.  
XX Liu J, Gurewicz V;

DR WPI; 1995-060991/08.

XX Pro-urokinase mutants - have thrombolytic activity but reduced  
PT fibrinogenolysis activity and non-specific plasminogen activation.  
XX Disclosure; Fig 1; 46pp; English.

XX AAR62991 is the wild type pro-urokinase, from which the new mutants  
CC described in AAR62992-R63008 were derived. These mutants retain the  
CC thrombolytic activity of the wild type protein, useful for the treatment  
CC of thromboembolism, but have a reduced fibrinogenolysis activity and non-  
CC specific plasminogen activation. The mutants can therefore be used for  
CC the lysis of fibrin clots without inducing systemic bleeding, as can be  
CC the case with the wild type protein. (Updated on 25-MAR-2003 to correct  
CC PN field.)  
XX Sequence 411 AA;

Query Match 100.0%; Score 1508; DB 2; Length 411;  
Best Local Similarity 100.0%; Pred. No. 2.8e-125;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0  
QY 1 KPSPPPEELKFCGQKTLPRPKLIIGGEFTTIENQFWFAAIYRRHGGSVTVVCGSLIS 60  
DB 136 KPSPPPEELKFCGQKTLPRPKLIIGGEFTTIENQFWFAAIYRRHGGSVTVVCGSLIS 195  
QY 61 PCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120  
DB 196 PCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 255  
QY 121 IALLKIRSKGRCAQPSRTIOTICLPMSYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT 180  
DB 256 IALLKIRSKGRCAQPSRTIOTICLPMSYNDPQFGTSCEITGFGKENSTDYLYPEQLKMT 315  
QY 181 VVKLISHRECOQPHYVGSVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240  
DB 316 VVKLISHRECOQPHYVGSVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 375  
QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276  
DB 376 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411

RESULT 40  
AAY39343

ID AAY39343 standard; protein; 411 AA.

XX AAY39343;  
DT 01-DEC-1999 (first entry)  
XX Human pro-urokinase.  
XX Serine protease; plasminogen; plasmin; activation; matrix; cancer;  
KW tumour; metastasis; X-ray crystallography; inhibitor.  
XX Homo sapiens.

XX Key Location/Qualifiers  
FT Protein 1..158  
FT /label= Mature\_urokinase\_A\_chain  
FT Protein 1..135  
FT /label= Low\_molecular\_weight\_urokinase  
FT Disulfide-bond 11..19  
FT Disulfide-bond 13..31  
FT Disulfide-bond 33..42  
FT Disulfide-bond 50..131  
FT Disulfide-bond 71..113  
FT Disulfide-bond 102..126  
FT Disulfide-bond 135..136  
FT Cleavage-site  
FT /note= "Additional cleavage at this site generates low  
FT molecular weight (LMW) urokinase"  
FT Disulfide-bond 148..279

FT FT /note= "Links mature urokinase A- and B-chains"  
FT Cleavage-site 159.159  
FT /note= "Cleavage at this site generates mature urokinase  
FT A- and B-chains"  
FT Protein 159.411  
FT /label= Mature\_urokinase\_B\_chain  
FT Disulfide-bond 189.205  
FT Disulfide-bond 197.268  
FT Disulfide-bond 293.362  
FT Modified-site 302  
FT /note= "N-glycosylated"  
FT Disulfide-bond 325.341  
FT Disulfide-bond 352.380  
FT Cleavage-site 405.406  
XX WO9945379-A2.  
XX  
XX  
XX 10-SEP-1999.  
XX  
XX 05-MAR-1999; 99WO-US004967.  
XX  
XX 06-MAR-1998; 98US-00036184.  
XX (ABBO ) ABBOTT LAB.  
XX Nienaber VL, Greer J, Abad-Zapatero C, Norbeck DW;  
XX WPI; 1999-571607/48.  
DR  
XX Identifying ligands for target biomolecules using X-ray crystallography.  
XX Example 1; Fig 5; 57pp; English.  
XX  
XX This sequence represents human pro-urokinase. The mature urokinase  
CC consists of an A- and B-chain, linked by a single disulphide bond, and is  
CC generated by proteolytic cleavage of the peptide bond between Lys 158 and  
CC Ile 159. Additional cleavage of the peptide bond between Lys 135 and Lys  
CC 136 generates a low molecular weight urokinase. The urokinase A-chain  
CC contains an EGF-like domain and a kringle domain, while the B-chain  
CC contains the catalytic domain. Urokinase is a serine protease and is  
CC strongly associated with tumour cells. Urokinase activates plasminogen  
CC which, in turn, activates the matrix metalloproteinases. Plasmin and the  
CC metalloproteinases degrade the extracellular matrix and promote tumour  
CC growth and metastasis. Inhibitors that specifically target urokinase may  
CC serve as effective anticancer agents. A novel method for identifying such  
CC ligands used X-ray crystallography to determine if a complex is formed  
CC between a ligand and a target biomolecule. However, crystals of a native  
CC urokinase/inhibitor complex had poor diffraction quality. Human urokinase  
CC was therefore engineered so that it would produce crystals with the  
CC desired qualities. This engineered urokinase was designated mu-UK  
XX (AA939344)  
XX  
XX Sequence 411 AA;  
SQ  
Query Match 100.0%; Score 1508; DB 2; Length 411;  
Best Local Similarity 100.0%; Pred. No. 2.8e-125;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KPSPPELKFQCCQKTLRPFKIIIGGEFTTIENQPFPAALYRHRGGSVTVYCGGSLIS 60  
DB 136 KPSPPELKFQCCQKTLRPFKIIIGGEFTTIENQPFPAALYRHRGGSVTVYCGGSLIS 195  
QY 61 PCWISATHCFIDYPKEDYIVYLGSRNLNNTQGMKFEVENILHKDYSADTLAHND 120  
DB 196 PCWISATHCFIDYPKEDYIVYLGSRNLNNTQGMKFEVENILHKDYSADTLAHND 255  
QY 121 IALLKIRSEGRCAQPSRTIOTICLPMSYNDPQGTSCITGFGKENSTDYLYPEQLKMT 180  
DB 256 IALLKIRSEGRCAQPSRTIOTICLPMSYNDPQGTSCITGFGKENSTDYLYPEQLKMT 315  
QY 181 VVKLISHRECCQPHYGSEVTTKMLCAADPWKMTSCQDSCGGPLVCSLQGRMTLTGIVS 240  
DB 316 VVKLISHRECCQPHYGSEVTTKMLCAADPWKMTSCQDSCGGPLVCSLQGRMTLTGIVS 375

QY 241 WGRGCAKDKPGVTVRVSHFLPWIRSHTKBENGIAL 276  
DB 376 WGRGCAKDKPGVTVRVSHFLPWIRSHTKBENGIAL 411  
RESULT 41  
AA42284  
ID AA42284 standard; protein; 411 AA.  
XX  
XX AA42284;  
XX  
XX 01-DEC-1999 (first entry)  
XX Human pro-urokinase.  
XX  
XX Serine protease; plasminogen; plasmin; activation; matrix; cancer;  
XX tumour; metastasis; X-ray crystallography; inhibitor.  
XX  
XX Homo sapiens.  
XX  
XX Location/Qualifiers  
FH Key 1.158  
FT Protein /label= Mature\_urokinase\_A\_chain  
FT Protein 1.135  
FT /label= Low\_molecular\_weight\_urokinase  
FT Disulfide-bond 11.19  
FT Disulfide-bond 13.31  
FT Disulfide-bond 33.42  
FT Disulfide-bond 50.131  
FT Disulfide-bond 71.113  
FT Disulfide-bond 102.126  
FT Cleavage-site 135.136  
FT /note= "Additional cleavage at this site generates low  
FT molecular weight (LMW) urokinase"  
FT Disulfide-bond 148.279  
FT /note= "Links mature urokinase A- and B-chains"  
FT Cleavage-site 158.159  
FT /note= "Cleavage at this site generates mature urokinase  
FT A- and B-chains"  
FT Protein 159.411  
FT /label= Mature\_urokinase\_B\_chain  
FT Disulfide-bond 189.205  
FT Disulfide-bond 197.268  
FT Disulfide-bond 293.362  
FT Modified-site 302  
FT /note= "N-glycosylated"  
FT Disulfide-bond 325.341  
FT Disulfide-bond 352.380  
FT Cleavage-site 405.406  
XX WO9945389-A2.  
XX  
XX 10-SEP-1999.  
XX  
XX 01-MAR-1999; 99WO-US004518.  
XX  
XX 06-MAR-1998; 98US-00036184.  
XX (ABBO ) ABBOTT LAB.  
XX  
XX Nienaber VL, Greer J, Abad-Zapatero C, Norbeck DW;  
XX WPI; 1999-551079/46.  
DR  
XX Identifying ligands for target biomolecules using X-ray crystallography.  
XX used for designing ligands with improved biological activity for target  
XX receptor.  
XX  
XX Example 1; Fig 5; 57pp; English.  
XX  
XX This sequence represents human pro-urokinase. The mature urokinase  
CC consists of an A- and B-chain, linked by a single disulphide bond, and

generated by proteolytic cleavage of the peptide bond between Lys 158 and Lys 159. Additional cleavage of the peptide bond between Lys 135 and Lys 136 generates a low molecular weight urokinase. The urokinase A-chain contains an EGF-like domain and a kringle domain, while the B-chain contains the catalytic domain. Urokinase is a serine protease and is strongly associated with tumour cells. Urokinase activates plasminogen which, in turn, activates the matrix metalloproteinases. Plasmin and the metalloproteinases degrade the extracellular matrix and promote tumour growth and metastasis. Inhibitors that specifically target urokinase may serve as effective anticancer agents. A novel method for identifying such ligands used X-ray crystallography to determine if a complex is formed between a ligand and a target biomolecule. However, crystals of a native urokinase/inhibitor complex had poor diffraction quality. Human urokinase was therefore engineered so that it would produce crystals with the desired qualities. This engineered urokinase was designated mu-UK (AA42285).

XX SQ Sequence 411 AA;

Query Match 100.0%; Score 1508; DB 2; Length 411;  
Best Local Similarity 100.0%; Pred. No. 2.8e-125;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCGQKTLRPFKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGSLIS 60  
DB 136 KPSSPPEELKFCGQKTLRPFKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGSLIS 195  
QY 61 PCWVISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKPEVENLILHKDYSADTLAHND 120  
DB 196 PCWVISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKPEVENLILHKDYSADTLAHND 255  
QY 121 IALLKIRSEKGRCAQPSRTIQTICLPWYNDPQFGTSCITGFGKENSVDLYPEQLKMT 180  
DB 256 IALLKIRSEKGRCAQPSRTIQTICLPWYNDPQFGTSCITGFGKENSVDLYPEQLKMT 315  
QY 181 VKVLSHRECCQPHYGVSEVTTKMLCAADPQWTKDSCQDSDGSLVCSLQGRMTLTGIS 240  
DB 316 VKVLSHRECCQPHYGVSEVTTKMLCAADPQWTKDSCQDSDGSLVCSLQGRMTLTGIS 375  
QY 241 WGRGCALDKDPGVYTRVSHFLPWIRSHKTEENGLAL 276  
DB 376 WGRGCALDKDPGVYTRVSHFLPWIRSHKTEENGLAL 411

RESULT 42

AA422836  
ID AA422836 standard; protein; 411 AA.

XX AC AA422836;

XX DT 29-AUG-2000 (first entry)

XX DE Urokinase plasminogen activator (uPA).

XX KW N-terminal; pro-uPA; urokinase plasminogen activator; receptor; uPAR;  
KW anti-cancer; anti-metastatic; anti-proliferative; anti-atherosclerotic;  
KW anti-thrombotic; anti-angiogenic; anti-inflammatory; anti-arthritis;  
KW anti-fibrotic; apoptotic; vasotropic; anti-diabetic; ophthalmological;  
XX thrombolytic.

XX OS Homo sapiens.

XX PH Key Location/Qualifiers

FT Disulfide-bond 11..19  
FT Disulfide-bond 13..31  
FT Disulfide-bond 33..42  
FT Disulfide-bond 50..131  
FT Disulfide-bond 71..113  
FT Disulfide-bond 102..126  
FT Disulfide-bond 148..279  
FT Disulfide-bond 189..205  
FT Disulfide-bond 197..268  
FT Disulfide-bond 293..362

FT Disulfide-bond 325..341  
FT Disulfide-bond 352..380  
XX W0200026353-A1.

XX PD 11-MAY-2000.

XX PF 28-OCT-1999; 99WO-US025210.

XX PR 29-OCT-1999; 98US-00181816.

XX PA (ANGS-) ANGSTROM PHARM INC.

XX PI Mazar AP, Jones TR;

XX DR WPI; 2000-365605/31.

XX PT New cyclic peptide, useful for treatment or diagnosis of e.g. tumors and other diseases involving cell proliferation or migration, targets the urokinase plasminogen activator receptor.

XX ES Disclosure; Fig 1; 93pp; English.

XX CC The present sequence shows the wild-type urokinase plasminogen activator (uPA). Cyclic peptides based on the amino acids residues 20-30 (the receptor-binding region) of uPA are claimed. These cyclic peptides target the uPA receptor (uPAR), allowing therapeutic or diagnostic agents to be delivered to uPAR-expressing cells. The cyclic peptides are used, optionally when linked to a therapeutic agent, to inhibit migration, invasion and proliferation of cells, or angiogenesis, or to induce apoptosis. Particularly they are used, in human or veterinary medicine, to treat diseases characterized by these processes, e.g. solid tumors, leukaemia or lymphoma (or their metastases); benign hyperplasia; atherosclerosis; restenosis; ischaemia; deep vein thrombosis; neovascular glaucoma; diabetic retinopathy; arthritis; fibrosis; bone fracture etc., most particularly growth, invasion and metastasis of tumors. When labeled, the cyclic peptides can be used for diagnostic detection of uPA (a marker of metastasis) on cells, tissues etc., in vivo or in vitro, an when immobilized they are used to isolate uPAR or cells that express them. The cyclic peptides are stable, soluble in water, bind strongly to uPAR, are relatively inexpensive to produce and may be derivatized by attachment of therapeutic or diagnostic agents without significantly affecting their binding. Since they target uPAR, they should have relatively low systemic toxicity and only low doses are required

XX SQ Sequence 411 AA;

Query Match 100.0%; Score 1508; DB 3; Length 411;

Best Local Similarity 100.0%; Pred. No. 2.8e-125;

Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps

QY 1 KPSSPPEELKFCGQKTLRPFKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGSLIS 6  
DB 136 KPSSPPEELKFCGQKTLRPFKIIGGEFTTIENQWFAAIYRRHGGSVTVVCGSLIS 1  
QY 61 PCWVISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKPEVENLILHKDYSADTLAHND 1  
DB 196 PCWVISATHCFIDYPKEDYIVYLGSRSLNSNTQGMKPEVENLILHKDYSADTLAHND 2  
QY 121 IALLKIRSEKGRCAQPSRTIQTICLPWYNDPQFGTSCITGFGKENSVDLYPEQLKMT 3  
DB 256 IALLKIRSEKGRCAQPSRTIQTICLPWYNDPQFGTSCITGFGKENSVDLYPEQLKMT 4  
QY 181 VKVLSHRECCQPHYGVSEVTTKMLCAADPQWTKDSCQDSDGSLVCSLQGRMTLTGIS 5  
DB 316 VKVLSHRECCQPHYGVSEVTTKMLCAADPQWTKDSCQDSDGSLVCSLQGRMTLTGIS 6  
QY 241 WGRGCALDKDPGVYTRVSHFLPWIRSHKTEENGLAL 276  
DB 376 WGRGCALDKDPGVYTRVSHFLPWIRSHKTEENGLAL 411

RESULT 43

AAB20489  
 ID AAB20489 standard; protein; 411 AA.  
 AC AAB20489;  
 XX  
 DT 21-JUN-2001 (first entry)  
 XX  
 DE Human pro-urokinase plasminogen activator.  
 XX  
 DE Urokinase plasminogen activator; uPA; human; tumour; cell migration;  
 KW cell invasion; cell proliferation; angiogenesis; apoptosis; antitumour;  
 KW diagnosis; therapy.  
 XX  
 OS Homo sapiens.  
 FH  
 FH Key Location/Qualifiers  
 FT Domain 1..135  
 FT /label= ATP  
 FT /note= "the ATF domain alternatively comprises residues 1  
 FT -43"  
 FT Domain 4..43  
 FT /label= Growth\_factor\_domain  
 FT Disulfide-bond 11..19  
 FT Disulfide-bond 13..31  
 FT Disulfide-bond 33..42  
 FT Disulfide-bond 50..131  
 FT Disulfide-bond 71..113  
 FT Disulfide-bond 102..126  
 FT Disulfide-bond 148..279  
 FT Cleavage-site 158..159  
 FT /note= "Cleavage at this site results in the formation of  
 FT the two-chain active uPA (tcuPA)"  
 FT Disulfide-bond 189..205  
 FT Disulfide-bond 197..268  
 FT Disulfide-bond 293..362  
 FT Disulfide-bond 325..341  
 FT Disulfide-bond 352..380  
 XX  
 PN WO200125410-A2.  
 XX  
 XX 12-APR-2001.  
 XX  
 PF 27-SEP-2000; 2000WO-US026502.  
 XX  
 PR 01-OCT-1999; 99US-0157012P.  
 XX  
 PA (ANGS-) ANGSTROM PHARM INC.  
 XX  
 PI Mazar AP, Jones TR;  
 XX  
 XX WPI; 2001-290611/30.  
 XX  
 XX Novel urokinase plasminogen activator cell surface receptor-targeting  
 PT protein or peptide, useful for inhibiting angiogenesis or cell migration,  
 PT invasion or proliferation, is diagnostically or therapeutically labeled.  
 XX  
 PS Disclosure; Fig 1; 35pp; English.  
 XX  
 XX The present sequence is that of human pro-urokinase plasminogen activator  
 CC (pro-uPA). The invention provides a uPA receptor (uPAR) targeting protein  
 CC or peptide that is labelled and used in methods of diagnosis and therapy.  
 CC The labelled protein or peptide preferably has the following properties:  
 CC it comprises at least 38 amino acid residues, including residues 13-30 of  
 CC the uPAR binding site of uPA; competes with labelled DRP-uPA for binding  
 CC to a cell or molecule that has a binding site for uPA; has an IC50 value  
 CC of about 10 nM or less; and is not a fusion protein. Preferred molecules  
 CC are uPA, (residues 1-411), single chain uPA, tcuPA (inactivated with the  
 CC suicide inhibitor diisopropyl fluorophosphate), the N-terminal ATF  
 CC fragment (amino acids 1-135 or 1-143) of uPA, or the growth factor domain  
 CC (residues 4-43). Suitable labels include a radionuclide, a PBT-imagable  
 CC agent, an MRI-imagable agent, a fluoroscer, a fluorogen, a chromophore,  
 CC a chromogen, a phosphorescer, a chemiluminescer or a bioluminescer. The  
 CC methods are used to inhibit cell migration, cell invasion (preferably

CC invasiveness of tumour cells), cell proliferation or angiogenesis, or to  
 CC induce apoptosis, preferably in the treatment of a subject having a  
 CC disease or condition associated with undesired cell migration, invasion,  
 CC proliferation or angiogenesis (claimed). The protein or peptide is also  
 CC useful for treating diseases or conditions including primary growth of a  
 CC solid tumour, leukaemia or lymphoma, tumour invasion, metastasis;  
 CC atherosclerosis, myocardial angiogenesis, telangiectasia, corneal  
 CC disease, rubeosis, neovascular glaucoma, diabetic and other retinopathy,  
 CC macular degeneration, arthritis, fibrosis, wound healing with scarring  
 CC and fibrosis, peptic ulcers, bone fracture, keloids, or a disorder of  
 CC vasculogenesis, haematopoiesis, ovulation, menstruation, pregnancy or  
 CC placental associated with pathogenic cell invasion or with  
 CC angiogenesis. The protein or peptide probe is internalised by the cells  
 CC to which it binds, e.g. tumour cells, and is useful for imaging  
 CC techniques in which it reduces the background signal relative to  
 CC specifically bound probes. This uptake permits clearance of circulating  
 CC probe so that the ratio of labelled probe inside tumour cells to the  
 CC probe elsewhere in the body increases  
 XX  
 SQ Sequence 411 AA;  
 Query Match 100.0%; Score 1508; DB 4; Length 411;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-125;  
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0  
 QY 1 KPSSPEELKFCQCKTLRPRFKIIGGEFTTIENQWFAIYRRHGGSVYVCGGSLIS 60  
 DB 136 KPSSPEELKFCQCKTLRPRFKIIGGEFTTIENQWFAIYRRHGGSVYVCGGSLIS 195  
 QY 61 PCWVISATHCFIDYPKKEDYVYLGSRSLNSNTQGEKFEVENLILHKQYSAADTLAHND 120  
 DB 196 PCWVISATHCFIDYPKKEDYVYLGSRSLNSNTQGEKFEVENLILHKQYSAADTLAHND 255  
 QY 121 IALLKIRSKGRCAQPSRTIQTCLPSWYNDPFGTSCETTGKENSITDLYPEOLKMT 180  
 DB 256 IALLKIRSKGRCAQPSRTIQTCLPSWYNDPFGTSCETTGKENSITDLYPEOLKMT 315  
 QY 181 VKLISHRECCQPHYVGSEVTTKMLCAADPQWKTDCQDGGPLVCSLQGRMTLGIVS 240  
 DB 316 VKLISHRECCQPHYVGSEVTTKMLCAADPQWKTDCQDGGPLVCSLQGRMTLGIVS 375  
 QY 241 WGRGKALKDKPGYTVRVSHPLPWIRSHKTEENGLAL 276  
 DB 376 WGRGKALKDKPGYTVRVSHPLPWIRSHKTEENGLAL 411  
 RESULT 44  
 AAB74797  
 ID AAB74797 standard; protein; 411 AA.  
 AC AAB74797;  
 XX  
 XX 12-JUN-2001 (first entry)  
 DT  
 DE Prourokinase protein sequence.  
 XX  
 XX Prourokinase; Pro-309; mutagenic; urokinase; zymogen; mutant;  
 KW lowered fibrinogen dissolving activity; fibrin; E segment; D segment;  
 KW lowered non-specific fibrin dissolving zymogen activation.  
 XX  
 OS Unidentified.  
 FH  
 FH Key Location/Qualifiers  
 FT Disulfide-bond 11..19  
 FT Disulfide-bond 13..31  
 FT Disulfide-bond 33..42  
 FT Disulfide-bond 50..131  
 FT Disulfide-bond 71..113  
 FT Disulfide-bond 102..126  
 FT Disulfide-bond 148..279  
 FT Disulfide-bond 189..205  
 FT Disulfide-bond 197..268  
 FT Disulfide-bond 293..362



FT Disulfide-bond 325. .341  
 FT Disulfide-bond 352. .380  
 XX CN1277262-A.  
 XX 20-DEC-2000.  
 XX 10-JUL-2000; 2000CN-00109829.  
 XX 10-JUL-2000; 2000CN-00109829.  
 XX (LIUJ/) LIU J.  
 XX Sun Z, Liu J;  
 XX WPI; 2001-266614/28.  
 XX Urokinase zymogen mutant.  
 XX Example; Fig 1; 11pp; Chinese.  
 XX The present invention describes a prourokinase mutant comprising the  
 CC amino acid sequence point mutation at proline 309. The mutation makes the  
 CC mutant have an intrinsic activity 2.5-20 times lower than that of natural  
 CC prourokinase, including lowered fibrinogen dissolving activity and  
 CC lowered non-specific fibrin dissolving zymogen activation. Compared with  
 CC the natural prourokinase, the fibrin dissolving zymogen activation of the  
 CC mutant may be promoted by not only the E segment of degraded fibrin but  
 CC also the D segment. The present sequence represents a wild type  
 CC prourokinase protein sequence which is used in an example from the  
 CC present invention. N.B. The sequence in the specification is of poor  
 CC quality so the sequence given here is of the indexers best interpretation  
 XX Sequence 411 AA;  
 XX  
 Query Match 100.0%; Score 1508; DB 4; Length 411;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-125;  
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KPSSPPEELKFCQCGQKTLRPRFKIIGGFTTIENQPFWFAIYRRHRGGSVTVVCGSLIS 60  
 DB 136 KPSSPPEELKFCQCGQKTLRPRFKIIGGFTTIENQPFWFAIYRRHRGGSVTVVCGSLIS 195  
 QY 61 PCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKEFEVENLIHKDYSADTLAHND 120  
 DB 196 PCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKEFEVENLIHKDYSADTLAHND 255  
 QY 121 IALLKIRSGRCAQPSRTIQTICLPMSYNDPQGTSCETITGFKENSTDYLYPEQLKMT 180  
 DB 256 IALLKIRSGRCAQPSRTIQTICLPMSYNDPQGTSCETITGFKENSTDYLYPEQLKMT 315  
 QY 181 VVKLISHRECCQPHYVGSEVTTKMLCAADPQWKTDCQDSDGGLVCSLQGRMTLTGIVS 240  
 DB 316 VVKLISHRECCQPHYVGSEVTTKMLCAADPQWKTDCQDSDGGLVCSLQGRMTLTGIVS 375  
 XX  
 RESULT 45  
 AAE16544  
 ID AAE16544 standard; protein; 411 AA.  
 XX  
 AC AAE16544;  
 XX  
 XX 09-APR-2002 (first entry)  
 XX Human urokinase-type plasminogen activator tcuPA and scuPA protein.  
 XX Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;  
 KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;  
 KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;

KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;  
 KW clotting disorder; uterine contraction disorder; respiratory disease;  
 KW male impotence; adult respiratory distress syndrome; tcuPA; scuPA;  
 XX two chain urokinase; single chain urokinase.  
 XX Homo sapiens.  
 XX OS  
 XX WO200197752-A2.  
 XX 27-DEC-2001.  
 XX 13-JUN-2001; 2001WO-US018976.  
 XX 20-JUN-2000; 2000US-0212874P.  
 XX (UYPE-) UNIV PENNSYLVANIA.  
 XX Cines DB, Higazi AA;  
 XX WPI; 2002-122240/16.  
 XX DR N-PSDB; AAD27077.  
 XX Composition for modulating muscle cell and tissue contractility for  
 PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,  
 PT comprising domains from urokinase-type plasminogen activator.  
 XX Claim 9; Fig 1C; 117pp; English.  
 XX The invention relates to a composition comprising one or more domains of  
 CC urokinase-type plasminogen activator (uPA). The composition is used to  
 CC modulate the contractility and angiogenic activity of a mammalian muscle,  
 CC endothelial cell or tissue. The composition is used for treating stroke,  
 CC hypertension, hyperextension, atherosclerosis, heart attack, microvascular  
 CC occlusions, thrombotic microangiopathies, surgically induced thrombotic  
 CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell  
 CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,  
 CC diabetic retinopathy, wound healing, clotting disorder, uterine  
 CC contraction disorder, male impotence, respiratory disease or condition  
 CC such as asthma, adult respiratory distress syndrome, primary pulmonary  
 CC hypertension, microvascular thrombotic occlusion, and a disorder  
 CC associated with chronic intrapulmonary fibrin formation. The present  
 CC sequence is human urokinase-type plasminogen activator (uPA) two chain  
 CC urokinase (tcuPA) and single chain urokinase (scuPA) protein  
 XX Sequence 411 AA;  
 XX  
 Query Match 100.0%; Score 1508; DB 5; Length 411;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-125;  
 Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0  
 QY 1 KPSSPPEELKFCQCGQKTLRPRFKIIGGFTTIENQPFWFAIYRRHRGGSVTVVCGSLIS 60  
 DB 136 KPSSPPEELKFCQCGQKTLRPRFKIIGGFTTIENQPFWFAIYRRHRGGSVTVVCGSLIS 195  
 QY 61 PCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKEFEVENLIHKDYSADTLAHND 120  
 DB 196 PCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKEFEVENLIHKDYSADTLAHND 255  
 QY 121 IALLKIRSGRCAQPSRTIQTICLPMSYNDPQGTSCETITGFKENSTDYLYPEQLKMT 180  
 DB 256 IALLKIRSGRCAQPSRTIQTICLPMSYNDPQGTSCETITGFKENSTDYLYPEQLKMT 315  
 QY 181 VVKLISHRECCQPHYVGSEVTTKMLCAADPQWKTDCQDSDGGLVCSLQGRMTLTGIVS 240  
 DB 316 VVKLISHRECCQPHYVGSEVTTKMLCAADPQWKTDCQDSDGGLVCSLQGRMTLTGIVS 375  
 XX  
 RESULT 46  
 ADE85977



```

QY 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
DB 386 WGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 421

RESULT 48
AAP50114
ID AAP50114 standard; protein; 431 AA.
XX AAP50114;
XX
DT 27-SEP-1991 (first entry)
DE Sequence encoded by the signal sequence and noncoding region of the pro-
DE UK structural gene (Sequence II).
XX
KW Enzyme; thrombosis therapy; embolic disease; single-chain pro-urokinase.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..20
FT Domain /label= signal peptide
FT Domain 21..177
FT Domain /label= A chain
FT Domain 178..431
FT Domain /label= B chain
XX
PN EP154272-A.
XX
PD 11-SEP-1985.
XX
PF 23-FEB-1985; 85EP-00102031.
XX
PR 27-FEB-1984; 84JP-00037119.
PR 31-JAN-1985; 85JP-00017969.
XX
PA (GREC ) GREEN CROSS CORP.
XX
PI Hiramatsu R, Kaneda T, Nagai M, Rimura H, Nishida M, Suyama T;
XX WPI; 1985-224693/37.
XX N-PSDB; AAN50138.
XX
PT Glycosylated single-chain pro-urokinase - prepd. by cultivating animal
PT cells transformed by DNA prepd. from m RNA.
XX
PS Disclosure; Page 8-10; 64pp; English.
XX
CC The inventors claim a method of producing single-chain pro-urokinase by
CC using as template, mRNA obtd. from cells of an established human kidney-
CC derived cell line. The urokinase is used to treat thrombosis and embolic
CC diseases as well as in the treatment of diseases in combination with
CC anticancer agents
XX
SQ Sequence 431 AA;

Query Match 100.0%; Score 1508; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 3e-125;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPBELKFCQCKTLRPFKIIIGGFTTIENQPFAAIYRRHGGSVYVCGSLIS 60
DB 156 KPSSPPBELKFCQCKTLRPFKIIIGGFTTIENQPFAAIYRRHGGSVYVCGSLIS 215

QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRRLNSNTGEMKFEVENLILHKDYSADTLAHHND 120
DB 216 PCWVISATHCFIDYPKKEDYIVYLGSRRLNSNTGEMKFEVENLILHKDYSADTLAHHND 275

QY 121 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCITGFGKENS TDLYLPEQLKWT 180
DB 276 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCITGFGKENS TDLYLPEQLKWT 335

us-09-880-503-5.rag
QY 181 VVKLISHRECQPHYIGSEVTTMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIYS 240
DB 336 VVKLISHRECQPHYIGSEVTTMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIYS 395

RESULT 49
AAP60783
ID AAP60783 standard; protein; 431 AA.
XX AAP60783;
XX
DT 25-MAR-2003 (revised)
DT 23-OCT-1991 (first entry)
XX
DE Human urokinase.
XX
KW E.coli; high molecular urokinase.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Protein 21..431
XX
PN JP61181377-A.
XX
PD 14-AUG-1986.
XX
PF 25-JAN-1985; 85JP-00011032.
XX
PR 25-JAN-1985; 85JP-00011032.
XX
PA (NISC ) NISSAN CHEM IND LTD.
PA (HODO ) HODOGAYA CHEM IND CO LTD.
PA (SAGA ) SAGAMI CHEM RES CENTRE.
PA (CENG ) CENTRAL GLASS CO LTD.
PA (NIPS ) NIPPON SODA CO.
PA (TOYJ ) TOYO SODA MFG CO LTD.
XX
XX WPI; 1986-254744/39.
XX N-PSDB; AAN60703.
XX
PT Human urokinase gene - has N-end of aminoacid sequence coded by codon
PT used in Escherichia coli.
XX
PS Disclosure; Fig 2; 19pp; Japanese.
XX
CC The claimed gene product may be expressed in a transformed E.coli host,
CC for the efficient production of high molecular human urokinase. The N-
CC terminal of the protein expressed by the transforming plasmid is replaced
CC with a codon frequently used in E.coli. (Updated on 25-MAR-2003 to
CC correct PA field.)
XX
SQ Sequence 431 AA;

Query Match 100.0%; Score 1508; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 3e-125;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPBELKFCQCKTLRPFKIIIGGFTTIENQPFAAIYRRHGGSVYVCGSLIS 60
DB 156 KPSSPPBELKFCQCKTLRPFKIIIGGFTTIENQPFAAIYRRHGGSVYVCGSLIS 215

QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRRLNSNTGEMKFEVENLILHKDYSADTLAHHND 120
DB 216 PCWVISATHCFIDYPKKEDYIVYLGSRRLNSNTGEMKFEVENLILHKDYSADTLAHHND 275

QY 121 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCITGFGKENS TDLYLPEQLKWT 180
DB 276 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCITGFGKENS TDLYLPEQLKWT 335

```



Sequence 594, Appl  
Sequence 4, Appl  
Sequence 1, Appl  
Sequence 562, Appl  
Sequence 6, Appl  
Sequence 3, Appl  
Sequence 266, Appl  
Sequence 2, Appl  
Sequence 47, Appl  
Sequence 591, Appl  
Sequence 6266, Appl  
Sequence 2927, Appl  
Sequence 39, Appl  
Sequence 17, Appl  
Sequence 18, Appl  
Sequence 16, Appl  
Sequence 11, Appl  
Sequence 12, Appl  
Sequence 13, Appl  
Sequence 14, Appl  
Sequence 12, Appl  
Sequence 13, Appl  
Sequence 14, Appl  
Sequence 15, Appl  
Sequence 15, Appl  
Sequence 16, Appl  
Sequence 10, Appl  
Sequence 11, Appl  
Sequence 3, Appl  
Sequence 3, Appl  
Sequence 3, Appl  
Sequence 18, Appl  
Sequence 19, Appl  
Sequence 1, Appl  
Sequence 203, Appl  
Sequence 7, Appl  
Sequence 145, Appl  
Sequence 26, Appl  
Sequence 8, Appl  
Sequence 4, Appl  
Sequence 26, Appl  
Sequence 26, Appl  
Sequence 1, Appl  
Sequence 1, Appl  
Sequence 46, Appl  
Sequence 2, Appl  
Sequence 30, Appl  
Sequence 2, Appl  
Sequence 28, Appl  
Sequence 72, Appl  
Sequence 26, Appl  
Sequence 97, Appl  
Sequence 98, Appl  
Sequence 6, Appl  
Sequence 214, Appl  
Sequence 100, Appl  
Sequence 5, Appl  
Sequence 46, Appl

437 12 US-10-087-192-594  
431 14 US-10-193-656-4  
431 9 US-09-264-468B-1  
431 12 US-10-282-174-562  
403 9 US-09-880-503-6  
268 15 US-10-407-821-3  
445 15 US-10-360-101-266  
246 9 US-09-264-468B-2  
443 10 US-09-898-837A-47  
443 12 US-10-087-192-591  
443 12 US-10-106-698-6266  
337 15 US-10-264-049-2927  
337 15 US-09-997-003-39  
337 15 US-09-987-457-17  
268 10 US-09-987-455-18  
308 10 US-09-987-457-16  
308 10 US-09-987-455-17  
331 10 US-09-987-457-11  
331 10 US-09-987-455-12  
335 10 US-09-987-457-13  
335 10 US-09-987-455-14  
339 10 US-09-987-455-13  
339 10 US-09-987-455-13  
343 10 US-09-987-457-14  
343 10 US-09-987-457-15  
343 10 US-09-987-455-15  
343 10 US-09-987-455-15  
343 10 US-09-987-455-15  
343 10 US-09-987-457-10  
354 10 US-09-987-455-11  
372 9 US-09-084-491A-3  
372 13 US-10-103-704-3  
372 13 US-10-103-704-3  
372 13 US-09-987-455-8  
527 10 US-09-987-457-18  
527 10 US-09-987-455-19  
527 12 US-10-432-842-1  
527 15 US-10-360-101-203  
527 9 US-09-969-271-7  
562 9 US-09-974-298-145  
562 12 US-10-411-037-26  
562 12 US-10-411-026-26  
562 14 US-10-193-656-8  
562 14 US-10-443-701-4  
562 16 US-10-410-962-26  
562 16 US-10-411-049-26  
650 15 US-10-401-077-1  
253 16 US-10-600-187-8  
246 10 US-09-898-837A-46  
615 10 US-09-858-909-2  
615 14 US-10-172-712-30  
615 16 US-10-449-132-2  
655 14 US-10-172-712-28  
229 12 US-09-825-751A-72  
280 15 US-10-051-874-26  
280 16 US-10-451-168-97  
282 16 US-10-451-168-98  
822 9 US-09-147-947-6  
259 15 US-10-107-782-214  
259 16 US-10-038-854-100  
249 9 US-09-961-721-5  
249 14 US-10-170-789-46

## ALIGNMENTS

## RESULT 1

US-09-880-503-5

; Sequence 5, Application US/09880503

; Patent No. US20020131964A1

; GENERAL INFORMATION:

; APPLICANT: CINES, Douglas B

; APPLICANT: HIGAZI, Abd Al-Root

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2004, 14:53:05 ; Search time 50.3138 Seconds

(without alignments)

1530.046 Million cell updates/sec

Title: US-09-880-503-5

Perfect score: 1508

Sequence: 1 KPSPPELKFQCCQKTLRP.....VSHFLPWIRSHTEKENGAL 276

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1149313 seqs, 278921704 residues

Total number of hits satisfying chosen parameters: 1149313

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1508	100.0	276	9 US-09-880-503-5	Sequence 5, Appl
2	1508	100.0	323	9 US-09-880-503-7	Sequence 7, Appl
3	1508	100.0	411	9 US-09-880-503-3	Sequence 3, Appl
4	1508	100.0	411	15 US-10-407-821-2	Sequence 2, Appl
5	1508	100.0	431	12 US-10-411-037-34	Sequence 34, Appl
6	1508	100.0	431	12 US-10-411-026-34	Sequence 34, Appl
7	1508	100.0	431	13 US-10-076-421-2	Sequence 2, Appl
8	1508	100.0	431	14 US-10-171-311-184	Sequence 184, Appl
9	1508	100.0	431	14 US-10-301-822-161	Sequence 161, Appl
10	1508	100.0	431	14 US-10-247-671-149	Sequence 149, Appl
11	1508	100.0	431	14 US-10-131-985-21	Sequence 21, Appl
12	1508	100.0	431	15 US-10-295-027-414	Sequence 414, Appl
13	1508	100.0	431	15 US-10-295-027-414	Sequence 1275, Appl
14	1508	100.0	431	16 US-10-410-962-34	Sequence 34, Appl
15	1508	100.0	431	16 US-10-411-049-34	Sequence 34, Appl

; TITLE OF INVENTION: TISSUE CONTRACTABILITY  
; FILE REFERENCE: 9596-331  
; CURRENT APPLICATION NUMBER: US/09/880,503  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/212,847  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 276  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-880-503-5

Query Match 100.0%; Score 1508; DB 9; Length 276;  
Best Local Similarity 100.0%; Pred. No. 5.8e-146;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQCKTLRPRFKIIGGEFTTIENQPFALYYRHRGGSVTVVCGSLIS 60  
DB 1 KPSSPPEELKFCQCKTLRPRFKIIGGEFTTIENQPFALYYRHRGGSVTVVCGSLIS 60  
QY 61 PCWISATHCFIDYPKKEDYVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120  
DB 61 PCWISATHCFIDYPKKEDYVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120  
QY 121 IALLKIRSEKGRCAQPSRTIOTICLPMSYNDPQFTSCETITGFGKENSTDYLYPEQLKMT 180  
DB 121 IALLKIRSEKGRCAQPSRTIOTICLPMSYNDPQFTSCETITGFGKENSTDYLYPEQLKMT 180  
QY 181 VVKLISHRECCQPHYVYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240  
DB 181 VVKLISHRECCQPHYVYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240  
QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276  
DB 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276

RESULT 2  
US-09-880-503-7  
; Sequence 7, Application US/09880503  
; Patent No. US20020131964A1  
; GENERAL INFORMATION:  
; APPLICANT: CINES, Douglas B  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND  
; TITLE OF INVENTION: TISSUE CONTRACTABILITY  
; FILE REFERENCE: 9596-331  
; CURRENT APPLICATION NUMBER: US/09/880,503  
; CURRENT FILING DATE: 2001-06-13  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 323  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-880-503-7

Query Match 100.0%; Score 1508; DB 9; Length 323;  
Best Local Similarity 100.0%; Pred. No. 7.2e-146;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQCKTLRPRFKIIGGEFTTIENQPFALYYRHRGGSVTVVCGSLIS 60  
DB 48 KPSSPPEELKFCQCKTLRPRFKIIGGEFTTIENQPFALYYRHRGGSVTVVCGSLIS 107  
QY 61 PCWISATHCFIDYPKKEDYVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120  
DB 108 PCWISATHCFIDYPKKEDYVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 167

QY 121 IALLKIRSEKGRCAQPSRTIOTICLPMSYNDPQFTSCETITGFGKENSTDYLYPEQLKMT 180  
DB 168 IALLKIRSEKGRCAQPSRTIOTICLPMSYNDPQFTSCETITGFGKENSTDYLYPEQLKMT 227  
QY 181 VVKLISHRECCQPHYVYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240  
DB 228 VVKLISHRECCQPHYVYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 287  
QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276  
DB 288 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 323

RESULT 3  
US-09-880-503-3  
; Sequence 3, Application US/09880503  
; Patent No. US20020131964A1  
; GENERAL INFORMATION:  
; APPLICANT: CINES, Douglas B  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND  
; TITLE OF INVENTION: TISSUE CONTRACTABILITY  
; FILE REFERENCE: 9596-331  
; CURRENT APPLICATION NUMBER: US/09/880,503  
; CURRENT FILING DATE: 2001-06-13  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 411  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-880-503-3

Query Match 100.0%; Score 1508; DB 9; Length 411;  
Best Local Similarity 100.0%; Pred. No. 9.9e-146;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 KPSSPPEELKFCQCKTLRPRFKIIGGEFTTIENQPFALYYRHRGGSVTVVCGSLIS 60  
DB 136 KPSSPPEELKFCQCKTLRPRFKIIGGEFTTIENQPFALYYRHRGGSVTVVCGSLIS 195  
QY 61 PCWISATHCFIDYPKKEDYVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120  
DB 196 PCWISATHCFIDYPKKEDYVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 255  
QY 121 IALLKIRSEKGRCAQPSRTIOTICLPMSYNDPQFTSCETITGFGKENSTDYLYPEQLKMT 180  
DB 256 IALLKIRSEKGRCAQPSRTIOTICLPMSYNDPQFTSCETITGFGKENSTDYLYPEQLKMT 315  
QY 181 VVKLISHRECCQPHYVYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240  
DB 316 VVKLISHRECCQPHYVYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 375  
QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276  
DB 376 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411

RESULT 4  
US-10-407-821-2  
; Sequence 2, Application US/10407821  
; Publication No. US20030219386A1  
; GENERAL INFORMATION:  
; APPLICANT: IDELL, STEVEN  
; TITLE OF INVENTION: INTRAPLEURAL SINGLE-CHAIN UROKINASE ALONE OR COMPLEXED  
; TITLE OF INVENTION: TO ITS SOLUBLE RECEPTOR PROTECTS AGAINST PLEURAL  
; TITLE OF INVENTION: ADHESIONS IN TETRACYCLINE-INDUCED PLEURITIS IN RABBITS  
; FILE REFERENCE: UTSN:022US  
; CURRENT APPLICATION NUMBER: US/10/407,821  
; CURRENT FILING DATE: 2003-04-04  
; PRIOR APPLICATION NUMBER: 60/414,202

PRIOR FILING DATE: 2002-09-27  
PRIOR APPLICATION NUMBER: 60/370,466  
PRIOR FILING DATE: 2002-04-05  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 411  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-407-821-2

Query Match 100.0%; Score 1508; DB 15; Length 411;  
Best Local Similarity 100.0%; Pred. No. 9.9e-145;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQGQKTLRPFKIIIGGEFTTIENQPFPAIYRRHGGSVTVVCGSLIS 60  
DB 136 KPSSPPEELKFCQGQKTLRPFKIIIGGEFTTIENQPFPAIYRRHGGSVTVVCGSLIS 195

QY 61 PCWVISATHCFIDYPKEDYIVYLGSRSLNSTQEMKFEVENILHKDYSADTLAHND 120  
DB 196 PCWVISATHCFIDYPKEDYIVYLGSRSLNSTQEMKFEVENILHKDYSADTLAHND 255

QY 121 IALLKIRSEKGRCAQPSRTIOTICLPSMYNDPQFTSCBITGFKENSTDYLYPEQLKMT 180  
DB 256 IALLKIRSEKGRCAQPSRTIOTICLPSMYNDPQFTSCBITGFKENSTDYLYPEQLKMT 315

QY 181 VKLISHRECOQPHYGVSEVTTMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240  
DB 316 VKLISHRECOQPHYGVSEVTTMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 375

QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276  
DB 376 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411

RESULT 5  
US-10-411-037-34  
Sequence 34, Application US/10411037  
Publication No. US20040043446A1  
GENERAL INFORMATION:  
APPLICANT: Neose Technologies, Inc.  
APPLICANT: DeFrees, Shawn  
APPLICANT: Zopf, David  
APPLICANT: Bayer, Robert  
APPLICANT: Hakes, David  
APPLICANT: Chen, Xi  
APPLICANT: Bowe, Caryn  
TITLE OF INVENTION: ALPHA GALACTOSIDASE A: REMODELING AND GLYCOCONJUGATION OF ALPHA  
FILE REFERENCE: 040853-01-5082  
CURRENT APPLICATION NUMBER: US/10/411,037  
CURRENT FILING DATE: 2003-04-09  
PRIOR APPLICATION NUMBER: US 60/328,523  
PRIOR FILING DATE: 2001-10-10  
PRIOR APPLICATION NUMBER: US 60/344,692  
PRIOR FILING DATE: 2001-10-19  
PRIOR APPLICATION NUMBER: US 60/397,292  
PRIOR FILING DATE: 2002-06-07  
PRIOR APPLICATION NUMBER: US 60/391,777  
PRIOR FILING DATE: 2002-06-25  
PRIOR APPLICATION NUMBER: US 60/396,594  
PRIOR FILING DATE: 2002-07-17  
PRIOR APPLICATION NUMBER: US 60/404,249  
PRIOR FILING DATE: 2002-08-16  
PRIOR APPLICATION NUMBER: US 60/407,527  
NUMBER OF SEQ ID NOS: 75  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 34  
LENGTH: 431  
TYPE: PRT  
ORGANISM: Homo sapiens

US-10-411-037-34

Query Match 100.0%; Score 1508; DB 12; Length 431;  
Best Local Similarity 100.0%; Pred. No. 1e-145;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQGQKTLRPFKIIIGGEFTTIENQPFPAIYRRHGGSVTVVCGSLIS 60  
DB 156 KPSSPPEELKFCQGQKTLRPFKIIIGGEFTTIENQPFPAIYRRHGGSVTVVCGSLIS 215

QY 61 PCWVISATHCFIDYPKEDYIVYLGSRSLNSTQEMKFEVENILHKDYSADTLAHND 120  
DB 216 PCWVISATHCFIDYPKEDYIVYLGSRSLNSTQEMKFEVENILHKDYSADTLAHND 275

QY 121 IALLKIRSEKGRCAQPSRTIOTICLPSMYNDPQFTSCBITGFKENSTDYLYPEQLKMT 180  
DB 276 IALLKIRSEKGRCAQPSRTIOTICLPSMYNDPQFTSCBITGFKENSTDYLYPEQLKMT 335

QY 181 VKLISHRECOQPHYGVSEVTTMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240  
DB 336 VKLISHRECOQPHYGVSEVTTMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 395

QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276  
DB 396 WGRGCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 6  
US-10-411-026-34  
Sequence 34, Application US/10411026  
Publication No. US20040063911A1  
GENERAL INFORMATION:  
APPLICANT: Neose Technologies, Inc.  
APPLICANT: DeFrees, Shawn  
APPLICANT: Zopf, David  
APPLICANT: Bayer, Robert  
APPLICANT: Hakes, David  
APPLICANT: Chen, Xi  
TITLE OF INVENTION: PROTEIN REMODELING METHODS AND PROTEINS/PEPTIDES PRODUCED BY TH  
FILE REFERENCE: 040853-01-5053  
CURRENT APPLICATION NUMBER: US/10/411,026  
CURRENT FILING DATE: 2003-04-09  
PRIOR APPLICATION NUMBER: US 60/328,523  
PRIOR FILING DATE: 2001-10-10  
PRIOR APPLICATION NUMBER: US 60/344,692  
PRIOR FILING DATE: 2001-10-19  
PRIOR APPLICATION NUMBER: US 60/387,292  
PRIOR FILING DATE: 2002-06-07  
PRIOR APPLICATION NUMBER: US 60/391,777  
PRIOR FILING DATE: 2002-06-25  
PRIOR APPLICATION NUMBER: US 60/396,594  
PRIOR FILING DATE: 2002-07-17  
PRIOR APPLICATION NUMBER: US 60/404,249  
PRIOR FILING DATE: 2002-08-16  
PRIOR APPLICATION NUMBER: US 60/407,527  
NUMBER OF SEQ ID NOS: 75  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 34  
LENGTH: 431  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-411-026-34

Query Match 100.0%; Score 1508; DB 12; Length 431;  
Best Local Similarity 100.0%; Pred. No. 1e-145;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQGQKTLRPFKIIIGGEFTTIENQPFPAIYRRHGGSVTVVCGSLIS 60  
DB 156 KPSSPPEELKFCQGQKTLRPFKIIIGGEFTTIENQPFPAIYRRHGGSVTVVCGSLIS 215

QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120  
DB 216 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 275  
QY 121 IALLKIRSKGRCQAQPSRTIOTICLPSMYNDPQFGTSCETGFGKENSTDYLYPEQLKMT 180  
DB 276 IALLKIRSKGRCQAQPSRTIOTICLPSMYNDPQFGTSCETGFGKENSTDYLYPEQLKMT 335  
QY 181 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240  
DB 336 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 395  
QY 241 WGRGKALDKPGVYTRVSHFLPWIRSHTKENGLAL 276  
DB 396 WGRGKALDKPGVYTRVSHFLPWIRSHTKENGLAL 431  
RESULT 7  
US-10-076-421-2  
; Sequence 2, Application US/10076421  
; Publication No. US20020193304A1  
; GENERAL INFORMATION:  
; APPLICANT: WADA, MANABU  
; APPLICANT: WADA, NAKO  
; TITLE OF INVENTION: ANTI-HIV AGENTS  
; FILE REFERENCE: HAYAK-9  
; CURRENT APPLICATION NUMBER: US/10/076,421  
; CURRENT FILING DATE: 2002-05-17  
; PRIOR APPLICATION NUMBER: JP 2001-42655  
; PRIOR FILING DATE: 2001-02-20  
; PRIOR APPLICATION NUMBER: JP 2001-184284  
; PRIOR FILING DATE: 2001-06-19  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 431  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-076-421-2  
Query Match 100.0%; Score 1508; DB 13; Length 431;  
Best Local Similarity 100.0%; Pred. No. ie-145;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KPSPPPEELKFCQGQKTLRPRFKIIGGEFTTIENQWFAAIYRHRGGSVTVVCGSLIS 60  
DB 156 KPSPPPEELKFCQGQKTLRPRFKIIGGEFTTIENQWFAAIYRHRGGSVTVVCGSLIS 215  
QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120  
DB 216 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 275  
QY 121 IALLKIRSKGRCQAQPSRTIOTICLPSMYNDPQFGTSCETGFGKENSTDYLYPEQLKMT 180  
DB 276 IALLKIRSKGRCQAQPSRTIOTICLPSMYNDPQFGTSCETGFGKENSTDYLYPEQLKMT 335  
QY 181 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240  
DB 336 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 395  
QY 241 WGRGKALDKPGVYTRVSHFLPWIRSHTKENGLAL 276  
DB 396 WGRGKALDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 8  
US-10-171-311-184  
; Sequence 184, Application US/10171311  
; Publication No. US20030087270A1  
; GENERAL INFORMATION:  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Chen, Yan  
; APPLICANT: Zhao, Xumei

APPLICANT: Monahan, John  
APPLICANT: Kamatkar, Shubhangi  
APPLICANT: Giatt, Karen  
APPLICANT: Ganavarapu, Manjula  
APPLICANT: Hoerish, Sebastian  
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY  
TITLE OF INVENTION: OF CERVICAL CANCER  
FILE REFERENCE: MRI-035  
CURRENT APPLICATION NUMBER: US/10/171,311  
CURRENT FILING DATE: 2002-06-12  
PRIOR APPLICATION NUMBER: US 60/298,159  
PRIOR FILING DATE: 2001-06-13  
PRIOR APPLICATION NUMBER: US 60/298,155  
PRIOR FILING DATE: 2001-06-13  
PRIOR APPLICATION NUMBER: US 60/335,936  
PRIOR FILING DATE: 2001-11-14  
NUMBER OF SEQ ID NOS: 238  
SOFTWARE: PastSeq for Windows Version 4.0  
SEQ ID NO 184  
LENGTH: 431  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-171-311-184  
Query Match 100.0%; Score 1508; DB 14; Length 431;  
Best Local Similarity 100.0%; Pred. No. ie-145;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KPSPPPEELKFCQGQKTLRPRFKIIGGEFTTIENQWFAAIYRHRGGSVTVVCGSLIS 60  
DB 156 KPSPPPEELKFCQGQKTLRPRFKIIGGEFTTIENQWFAAIYRHRGGSVTVVCGSLIS 215  
QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120  
DB 216 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 275  
QY 121 IALLKIRSKGRCQAQPSRTIOTICLPSMYNDPQFGTSCETGFGKENSTDYLYPEQLKMT 180  
DB 276 IALLKIRSKGRCQAQPSRTIOTICLPSMYNDPQFGTSCETGFGKENSTDYLYPEQLKMT 335  
QY 181 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240  
DB 336 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 395  
QY 241 WGRGKALDKPGVYTRVSHFLPWIRSHTKENGLAL 276  
DB 396 WGRGKALDKPGVYTRVSHFLPWIRSHTKENGLAL 431  
RESULT 9  
US-10-301-822-161  
; Sequence 161, Application US/10301822  
; Publication No. US20030148410A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Berger, Allison  
; APPLICANT: Guillemette, Tracy L.  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Monahan, John E.  
; APPLICANT: Thibodeau, Stephen N.  
; APPLICANT: Burgart, Lawrence J.  
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AN  
THERAPY OF COLON CANCER  
FILE REFERENCE: MP01-029P2RNM  
CURRENT APPLICATION NUMBER: US/10/301,822  
CURRENT FILING DATE: 2002-11-21  
PRIOR APPLICATION NUMBER: US 60/339,971  
PRIOR FILING DATE: 2001-12-10  
PRIOR APPLICATION NUMBER: US 60/361,978  
PRIOR FILING DATE: 2002-03-05



```
; PRIOR APPLICATION NUMBER: US 60/381,988
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 161
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-301-822-161

Query Match      100.0%; Score 1508; DB 14; Length 431;
Best Local Similarity 100.0%; Pred. No. 1e-145;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQGOKTLRPRFKLIIGGEFTTIENQWFAAIYRRHGGSVTVYVCGSLIS 60
DB 156 KPSSPPEELKFCQGOKTLRPRFKLIIGGEFTTIENQWFAAIYRRHGGSVTVYVCGSLIS 215
QY 61 PCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
DB 216 PCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 275
QY 121 IALLKIRSKGRCAPSRITQICLPSMYNDPQFTSCETITGFGKENSTDYLYPEQLKMT 180
DB 276 IALLKIRSKGRCAPSRITQICLPSMYNDPQFTSCETITGFGKENSTDYLYPEQLKMT 335
QY 181 VVKLISHRECQOQPHYGGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
DB 336 VVKLISHRECQOQPHYGGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 395
QY 241 WGGCALCKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
DB 396 WGGCALCKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 11
US-10-131-985-21
; Sequence 21, Application US/10131985
; Publication No. US20030199440A1
; GENERAL INFORMATION:
; APPLICANT: Dack, Kevin N
; APPLICANT: Davies, Michael J
; APPLICANT: Fish, Paul V
; APPLICANT: Huggins, Jonathan P
; APPLICANT: McIntosh, Fraser S
; APPLICANT: Ocleston, Nicholas L
; TITLE OF INVENTION: Composition
; FILE REFERENCE: PCS 10391A
; CURRENT APPLICATION NUMBER: US/10/131,985
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: US/09/726,295
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: GB 9930768.8
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 21
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-985-21

Query Match      100.0%; Score 1508; DB 14; Length 431;
Best Local Similarity 100.0%; Pred. No. 1e-145;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQGOKTLRPRFKLIIGGEFTTIENQWFAAIYRRHGGSVTVYVCGSLIS 60
DB 156 KPSSPPEELKFCQGOKTLRPRFKLIIGGEFTTIENQWFAAIYRRHGGSVTVYVCGSLIS 215
QY 61 PCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
DB 216 PCWVISATHCFIDYPKKEDIYVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHND 275
QY 121 IALLKIRSKGRCAPSRITQICLPSMYNDPQFTSCETITGFGKENSTDYLYPEQLKMT 180
DB 276 IALLKIRSKGRCAPSRITQICLPSMYNDPQFTSCETITGFGKENSTDYLYPEQLKMT 335
QY 181 VVKLISHRECQOQPHYGGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240
DB 336 VVKLISHRECQOQPHYGGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 395
QY 241 WGGCALCKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
DB 396 WGGCALCKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 12
US-10-295-027-414
; Sequence 414, Application US/10295027
; Publication No. US2003023350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
```

APPLICANT: Gish, Kurt C.  
APPLICANT: Glynn, Richard  
APPLICANT: Hevezi, Peter A.  
APPLICANT: Mack, David H.  
APPLICANT: Murray, Richard  
APPLICANT: Watson, Susan R.  
APPLICANT: Eos Biotechnology, Inc.  
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and  
TITLE OF INVENTION: Methods of Screening for Modulators of Cancer  
FILE REFERENCE: 018501-012500US  
CURRENT APPLICATION NUMBER: US/10/295,027  
CURRENT FILING DATE: 2002-11-13  
PRIOR APPLICATION NUMBER: US 09/663,733  
PRIOR FILING DATE: 2000-09-15  
PRIOR APPLICATION NUMBER: US 60/350,666  
PRIOR FILING DATE: 2001-11-13  
PRIOR APPLICATION NUMBER: US 60/335,394  
PRIOR FILING DATE: 2001-11-15  
PRIOR APPLICATION NUMBER: US 60/332,464  
PRIOR FILING DATE: 2001-11-21  
PRIOR APPLICATION NUMBER: US 60/334,393  
PRIOR FILING DATE: 2001-11-29  
PRIOR APPLICATION NUMBER: US 60/340,376  
PRIOR FILING DATE: 2001-12-14  
PRIOR APPLICATION NUMBER: US 60/347,211  
PRIOR FILING DATE: 2002-01-08  
PRIOR APPLICATION NUMBER: US 60/347,349  
PRIOR FILING DATE: 2002-01-10  
PRIOR APPLICATION NUMBER: US 60/355,250  
PRIOR FILING DATE: 2002-02-08  
PRIOR APPLICATION NUMBER: US 60/356,714  
PRIOR FILING DATE: 2002-02-13  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 1386  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 414  
LENGTH: 431  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-295-027-414

Query Match 100.0%; Score 1508; DB 15; Length 431;  
Best Local Similarity 100.0%; Pred. No. 1e-145;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPFPAIYRHRGGSVTVVCGSLIS 60  
DB 156 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPFPAIYRHRGGSVTVVCGSLIS 215  
QY 61 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGENKFEVENLILHKDYSADTLAHHND 120  
DB 216 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGENKFEVENLILHKDYSADTLAHHND 275  
QY 121 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQGTSCITGPKENSTDYLYPEQLKMT 180  
DB 276 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQGTSCITGPKENSTDYLYPEQLKMT 335  
QY 181 VKLISHRECCQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240  
DB 336 VKLISHRECCQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 395  
QY 241 WGRGALKDKPGVTVRVSHFLPWIRSHTKENGIAL 276  
DB 396 WGRGALKDKPGVTVRVSHFLPWIRSHTKENGIAL 431

RESULT 13  
US-10-295-027-1275  
Sequence 1275, Application US/10295027  
Publication No. US20030232350A1  
GENERAL INFORMATION:  
APPLICANT: Afar, Daniel  
APPLICANT: Aziz, Natasha

APPLICANT: Ginsberg, Wendy M.  
APPLICANT: Gish, Kurt C.  
APPLICANT: Glynn, Richard  
APPLICANT: Hevezi, Peter A.  
APPLICANT: Mack, David H.  
APPLICANT: Murray, Richard  
APPLICANT: Watson, Susan R.  
APPLICANT: Eos Biotechnology, Inc.  
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and  
TITLE OF INVENTION: Methods of Screening for Modulators of Cancer  
FILE REFERENCE: 018501-012500US  
CURRENT APPLICATION NUMBER: US/10/295,027  
CURRENT FILING DATE: 2002-11-13  
PRIOR APPLICATION NUMBER: US 09/663,733  
PRIOR FILING DATE: 2000-09-15  
PRIOR APPLICATION NUMBER: US 60/350,666  
PRIOR FILING DATE: 2001-11-13  
PRIOR APPLICATION NUMBER: US 60/335,394  
PRIOR FILING DATE: 2001-11-15  
PRIOR APPLICATION NUMBER: US 60/332,464  
PRIOR FILING DATE: 2001-11-21  
PRIOR APPLICATION NUMBER: US 60/334,393  
PRIOR FILING DATE: 2001-11-29  
PRIOR APPLICATION NUMBER: US 60/340,376  
PRIOR FILING DATE: 2001-12-14  
PRIOR APPLICATION NUMBER: US 60/347,211  
PRIOR FILING DATE: 2002-01-08  
PRIOR APPLICATION NUMBER: US 60/347,349  
PRIOR FILING DATE: 2002-01-10  
PRIOR APPLICATION NUMBER: US 60/355,250  
PRIOR FILING DATE: 2002-02-08  
PRIOR APPLICATION NUMBER: US 60/356,714  
PRIOR FILING DATE: 2002-02-13  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 1386  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1275  
LENGTH: 431  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-295-027-1275

Query Match 100.0%; Score 1508; DB 15; Length 431;  
Best Local Similarity 100.0%; Pred. No. 1e-145;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPFPAIYRHRGGSVTVVCGSLIS 60  
DB 156 KPSSPPEELKFCQCKQLRPRFKIIGGEFTTIENQPFPAIYRHRGGSVTVVCGSLIS 215  
QY 61 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGENKFEVENLILHKDYSADTLAHHND 120  
DB 216 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGENKFEVENLILHKDYSADTLAHHND 275  
QY 121 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQGTSCITGPKENSTDYLYPEQLKMT 180  
DB 276 IALLKIRSKGRCAQPSRTIQTICLPSMYNDPQGTSCITGPKENSTDYLYPEQLKMT 335  
QY 181 VKLISHRECCQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240  
DB 336 VKLISHRECCQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 395  
QY 241 WGRGALKDKPGVTVRVSHFLPWIRSHTKENGIAL 276  
DB 396 WGRGALKDKPGVTVRVSHFLPWIRSHTKENGIAL 431

RESULT 14  
US-10-410-962-34  
Sequence 34, Application US/10410962  
Publication No. US20040077836A1  
GENERAL INFORMATION:  
APPLICANT: Neose Technologies, Inc.

APPLICANT: DeFrees, Shawn  
APPLICANT: Zopf, David  
APPLICANT: Bayer, Robert  
APPLICANT: Hakes, David  
APPLICANT: Chen, Xi  
APPLICANT: Bove, Caryn  
TITLE OF INVENTION: GRANULOCYTE COLONY STIMULATING FACTOR: REMODELING AND  
TITLE OF INVENTION: GLYCOCONJUGATION OF G-CSF  
FILE REFERENCE: 040853-01-5054  
CURRENT APPLICATION NUMBER: US/10/410,962  
CURRENT FILING DATE: 2003-04-09  
PRIOR APPLICATION NUMBER: US 60/328,523  
PRIOR FILING DATE: 2001-10-10  
PRIOR APPLICATION NUMBER: US 60/344,692  
PRIOR FILING DATE: 2001-10-19  
PRIOR APPLICATION NUMBER: US 60/387,292  
PRIOR FILING DATE: 2002-06-07  
PRIOR APPLICATION NUMBER: US 60/391,777  
PRIOR FILING DATE: 2002-06-25  
PRIOR APPLICATION NUMBER: US 60/396,594  
PRIOR FILING DATE: 2002-07-17  
PRIOR APPLICATION NUMBER: US 60/404,249  
PRIOR FILING DATE: 2002-08-16  
PRIOR APPLICATION NUMBER: US 60/407,527  
PRIOR FILING DATE: 2002-08-28  
NUMBER OF SEQ ID NOS: 75  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 34  
LENGTH: 431  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-410-962-34

Query Match 100.0%; Score 1508; DB 16; Length 431;  
Best Local Similarity 100.0%; Pred. No. 1e-145;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQCKTLRPRKLIIGGFTTIENQWFAAIYRRHGGSVTVVCGSLIS 60  
DB 156 KPSSPPEELKFCQCKTLRPRKLIIGGFTTIENQWFAAIYRRHGGSVTVVCGSLIS 215

QY 61 PCWVISATHCFDYPKKEDYIVYLGSRSLNSTQGMKFEVENLILHKDYSADTLAHND 120  
DB 216 PCWVISATHCFDYPKKEDYIVYLGSRSLNSTQGMKFEVENLILHKDYSADTLAHND 275

QY 121 IALLKTRSEKRCQAQPSRTIQTICLPSMYNDPQFQTSCEITGFGKENSTDYLYPEQLKWT 180  
DB 276 IALLKTRSEKRCQAQPSRTIQTICLPSMYNDPQFQTSCEITGFGKENSTDYLYPEQLKWT 335

QY 181 VVKLISHRECCQPHYVGSVTTKMLCAADPQWKTDCQDGGPLVCSLQGRMTLTGIVS 240  
DB 336 VVKLISHRECCQPHYVGSVTTKMLCAADPQWKTDCQDGGPLVCSLQGRMTLTGIVS 395

QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 276  
DB 396 WGRGCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431

RESULT 15  
US-10-411-049-34  
Sequence 34, Application US/10411049  
Publication No. US20040082026A1  
GENERAL INFORMATION:  
APPLICANT: Neose Technologies, Inc.  
APPLICANT: DeFrees, Shawn  
APPLICANT: Zopf, David  
APPLICANT: Bayer, Robert  
APPLICANT: Hakes, David  
APPLICANT: Chen, Xi  
APPLICANT: Bove, Caryn  
TITLE OF INVENTION: INTERFERON ALPHA: REMODELING AND GLYCOCONJUGATION OF INTERFERON  
TITLE OF INVENTION: ALPHA  
FILE REFERENCE: 040853-01-5055

CURRENT APPLICATION NUMBER: US/10/411,049  
CURRENT FILING DATE: 2003-04-09  
PRIOR APPLICATION NUMBER: US 60/328,523  
PRIOR FILING DATE: 2001-10-10  
PRIOR APPLICATION NUMBER: US 60/344,692  
PRIOR FILING DATE: 2001-10-19  
PRIOR APPLICATION NUMBER: US 60/387,292  
PRIOR FILING DATE: 2002-06-07  
PRIOR APPLICATION NUMBER: US 60/391,777  
PRIOR FILING DATE: 2002-06-25  
PRIOR APPLICATION NUMBER: US 60/396,594  
PRIOR FILING DATE: 2002-07-17  
PRIOR APPLICATION NUMBER: US 60/404,249  
PRIOR FILING DATE: 2002-08-16  
PRIOR APPLICATION NUMBER: US 60/407,527  
PRIOR FILING DATE: 2002-08-28  
NUMBER OF SEQ ID NOS: 75  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 34  
LENGTH: 431  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-411-049-34

Query Match 100.0%; Score 1508; DB 16; Length 431;  
Best Local Similarity 100.0%; Pred. No. 1e-145;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCQCKTLRPRKLIIGGFTTIENQWFAAIYRRHGGSVTVVCGSLIS 60  
DB 156 KPSSPPEELKFCQCKTLRPRKLIIGGFTTIENQWFAAIYRRHGGSVTVVCGSLIS 215

QY 61 PCWVISATHCFDYPKKEDYIVYLGSRSLNSTQGMKFEVENLILHKDYSADTLAHND 120  
DB 216 PCWVISATHCFDYPKKEDYIVYLGSRSLNSTQGMKFEVENLILHKDYSADTLAHND 275

QY 121 IALLKTRSEKRCQAQPSRTIQTICLPSMYNDPQFQTSCEITGFGKENSTDYLYPEQLKWT 180  
DB 276 IALLKTRSEKRCQAQPSRTIQTICLPSMYNDPQFQTSCEITGFGKENSTDYLYPEQLKWT 335

QY 181 VVKLISHRECCQPHYVGSVTTKMLCAADPQWKTDCQDGGPLVCSLQGRMTLTGIVS 240  
DB 336 VVKLISHRECCQPHYVGSVTTKMLCAADPQWKTDCQDGGPLVCSLQGRMTLTGIVS 395

QY 241 WGRGCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 276  
DB 396 WGRGCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431

RESULT 16  
US-10-087-192-594  
Sequence 594, Application US/10087192  
Publication No. US20020182586A1  
GENERAL INFORMATION:  
APPLICANT: Morris, David W.  
APPLICANT: Engelhard, Eric K.  
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
TITLE OF INVENTION: CANCER  
FILE REFERENCE: 529452000122  
CURRENT APPLICATION NUMBER: US/10/087,192  
CURRENT FILING DATE: 2002-03-01  
PRIOR APPLICATION NUMBER: US 09/747,377  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: US 09/798,586  
PRIOR FILING DATE: 2001-03-02  
NUMBER OF SEQ ID NOS: 2059  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 594  
LENGTH: 437  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-087-192-594

Query Match 100.0%; Score 1508; DB 12; Length 437;  
Best Local Similarity 100.0%; Pred. No. 1.1e-145;  
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPSSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRHRGGSVTVYCGGSLIS 60  
Db 162 KPSSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRHRGGSVTVYCGGSLIS 221

Qy 61 PCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 120  
Db 222 PCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 281

Qy 121 IALLKIRSKGRCACQPSRTIQTICLPSMYNDPQGTSCETITGFGKENSTDYLYPEQLKMT 180  
Db 282 IALLKIRSKGRCACQPSRTIQTICLPSMYNDPQGTSCETITGFGKENSTDYLYPEQLKMT 341

Qy 181 VVKLISHRECQOPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240  
Db 342 VVKLISHRECQOPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 401

Qy 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKBENGAL 276  
Db 402 WGRGALKDKPGVYTRVSHFLPWIRSHTKBENGAL 437

RESULT 17  
US-10-193-656-4  
; Sequence 4, Application US/10193656  
; Publication No. US20030096733A1  
; GENERAL INFORMATION:  
; APPLICANT: NY, Tor  
; APPLICANT: HOLMDEHL, Rikard  
; APPLICANT: LI, Jinan  
; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS  
; FILE REFERENCE: 3810/12577-US3  
; CURRENT APPLICATION NUMBER: US/10/193,656  
; CURRENT FILING DATE: 2002-07-10  
; PRIOR APPLICATION NUMBER: US 60/304,461  
; PRIOR FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: US 60/304,490  
; PRIOR FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: US 60/305,182  
; PRIOR FILING DATE: 2001-07-13  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 431  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; PUBLICATION INFORMATION:  
; DATABASE ACCESSION NUMBER: GenBank / P00749  
; DATABASE ENTRY DATE: 1986-07-21  
; RELEVANT RESIDUES: (1)..(431)  
US-10-193-656-4

Query Match 99.8%; Score 1505; DB 14; Length 431;  
Best Local Similarity 99.6%; Pred. No. 2.1e-145;  
Matches 275; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPSSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRHRGGSVTVYCGGSLIS 60  
Db 156 KPSSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRHRGGSVTVYCGGSLIS 215

Qy 61 PCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 120  
Db 216 PCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 275

Qy 121 IALLKIRSKGRCACQPSRTIQTICLPSMYNDPQGTSCETITGFGKENSTDYLYPEQLKMT 180  
Db 276 IALLKIRSKGRCACQPSRTIQTICLPSMYNDPQGTSCETITGFGKENSTDYLYPEQLKMT 335

Qy 181 VVKLISHRECQOPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240

Db 336 VVKLISHRECQOPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 395

Qy 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKBENGAL 276  
Db 396 WGRGALKDKPGVYTRVSHFLPWIRSHTKBENGAL 431

RESULT 18  
US-09-264-468B-1  
; Sequence 1, Application US/09264468B  
; Patent No. US20020106775A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Jieyi  
; APPLICANT: Nienaber, Vicki L.  
; APPLICANT: Henkin, Jack  
; APPLICANT: Smith, Richard A.  
; APPLICANT: Walter, Karl A.  
; APPLICANT: Severin, Jean M.  
; APPLICANT: Edalji, Rohinton  
; APPLICANT: Johnson Jr., Robert W.  
; APPLICANT: Holzman, Thomas F.  
; TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE  
; FILE REFERENCE: 6310.US.P1  
; CURRENT APPLICATION NUMBER: US/09/264,468B  
; CURRENT FILING DATE: 1999-03-05  
; PRIOR APPLICATION NUMBER: US 09/036,361  
; PRIOR FILING DATE: 1998-03-06  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 431  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: (1)..(20)  
; OTHER INFORMATION: Leader sequence  
; NAME/KEY: VARIANT  
; LOCATION: (279)..(279)  
; OTHER INFORMATION: Xaa = any amino acid  
; NAME/KEY: VARIANT  
; LOCATION: (302)..(302)  
; OTHER INFORMATION: Xaa = any amino acid  
US-09-264-468B-1

Query Match 98.8%; Score 1490; DB 9; Length 431;  
Best Local Similarity 99.3%; Pred. No. 7.3e-144;  
Matches 274; Conservative 0; Mismatches 2; Indels 0; Gaps 0

Qy 1 KPSSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRHRGGSVTVYCGGSLIS 60  
Db 156 KPSSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRHRGGSVTVYCGGSLIS 215

Qy 61 PCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 120  
Db 216 PCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHND 275

Qy 121 IALLKIRSKGRCACQPSRTIQTICLPSMYNDPQGTSCETITGFGKENSTDYLYPEQLKMT 180  
Db 276 IALLKIRSKGRCACQPSRTIQTICLPSMYNDPQGTSCETITGFGKENSTDYLYPEQLKMT 335

Qy 181 VVKLISHRECQOPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 240  
Db 336 VVKLISHRECQOPHYGSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQGRMTLTGIVS 395

Qy 241 WGRGALKDKPGVYTRVSHFLPWIRSHTKBENGAL 276  
Db 396 WGRGALKDKPGVYTRVSHFLPWIRSHTKBENGAL 431

RESULT 19  
US-10-282-174-562  
; Sequence 562, Application US/10282174

```

; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-09-880-503-6

```

```

Query Match 97.1%; Score 1465; DB 9; Length 403;
Best Local Similarity 99.6%; Pred. No. 2.5e-141;
Matches 268; Conservative 1; Mismatches 0; Indels 0; Gaps 0

QY 8 ELKFCGQKTLRPRKIIIGGFTTIENQWFAAIYRRHGGSVTVVCGSLISPCWVISA 67
DB 135 KLFQCGQKTLRPRKIIIGGFTTIENQWFAAIYRRHGGSVTVVCGSLISPCWVISA 194
QY 68 THCFIDYPKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHHNDIALKIR 127
DB 195 THCFIDYPKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHHNDIALKIR 254
QY 128 SKEGRCAQPSRTIQTICLPSMYNDPOFGTSCEITGKNSNDYLYPEQLKMTVVKLISH 187
DB 255 SKEGRCAQPSRTIQTICLPSMYNDPOFGTSCEITGKNSNDYLYPEQLKMTVVKLISH 314
QY 188 RECQPHYGVSEVTTKMLCAADPQWKTDCQDGGPLVCSLQGRMTLTGIVSMGRGAL 247
DB 315 RECQPHYGVSEVTTKMLCAADPQWKTDCQDGGPLVCSLQGRMTLTGIVSMGRGAL 374
QY 248 KDKPGVTVRVSHFLPWRSHRKEENGLAL 276
DB 375 KDKPGVTVRVSHFLPWRSHRKEENGLAL 403

```

```

RESULT 21
US-10-407-821-3
; Sequence 3, Application US/10407821
; Publication No. US20030219386A1
; GENERAL INFORMATION:
; APPLICANT: IDELL, STEVEN
; TITLE OF INVENTION: INTRAPLEURAL SINGLE-CHAIN UROKINASE ALONE OR COMPLEXED
; TITLE OF INVENTION: TO ITS SOLUBLE RECEPTOR PROTECTS AGAINST PLEURAL
; TITLE OF INVENTION: ADHESIONS IN TETRACYCLINE-INDUCED PLEURITIS IN RABBITS
; FILE REFERENCE: UTSN:02205
; CURRENT APPLICATION NUMBER: US/10/407,821
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 60/414,202
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 60/370,466
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-10-407-821-3

```

```

Query Match 97.1%; Score 1464; DB 15; Length 268;
Best Local Similarity 100.0%; Pred. No. 1.8e-141;
Matches 268; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 9 LKFCGQKTLRPRKIIIGGFTTIENQWFAAIYRRHGGSVTVVCGSLISPCWVISA 68
DB 1 LKFCGQKTLRPRKIIIGGFTTIENQWFAAIYRRHGGSVTVVCGSLISPCWVISA 60

```

```

; Publication No. US20030224380A1
; GENERAL INFORMATION:
; APPLICANT: Becker, Kenneth David
; APPLICANT: Velicelebi, Gonul
; APPLICANT: Elliot, Kathryn J.
; APPLICANT: Wang, Xin
; APPLICANT: Tanzi, Rudolph E.
; APPLICANT: Bertram, Lars
; APPLICANT: Saunders, Aleister J.
; APPLICANT: Mullin, Kristina M.
; APPLICANT: Sampson, Andrew Johnson
; APPLICANT: Blacker, Deborah Lynne
; TITLE OF INVENTION: GENES AND POLYMORPHISMS ON CHROMOSOME 10
; TITLE OF INVENTION: ASSOCIATED WITH ALZHEIMER'S DISEASE AND OTHER
; TITLE OF INVENTION: NEURODEGENERATIVE DISEASES
; FILE REFERENCE: 37481-3308
; CURRENT APPLICATION NUMBER: US/10/282,174
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: US 60/339,525
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 60/338,010
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/336,929
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/338,363
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 60/337,052
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: US 60/368,919
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 564
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 562
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
; NAME/KEY: VARIANT
; LOCATION: 15, 58, 141, 214, 231, 274, 366
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-282-174-562

```

```

Query Match 98.4%; Score 1484; DB 12; Length 431;
Best Local Similarity 98.6%; Pred. No. 3e-143;
Matches 272; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 KPSSPPEELKFCGQKTLRPRKIIIGGFTTIENQWFAAIYRRHGGSVTVVCGSLIS 60
DB 156 KPSSPPEELKFCGQKTLRPRKIIIGGFTTIENQWFAAIYRRHGGSVTVVCGSLIS 215
QY 61 PCWVISATHCFIDYPKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHHND 120
DB 216 PCWVISATHCFIDYPKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHHND 275
QY 121 IALLKIRSGRCAQPSRTIQTICLPSMYNDPOFGTSCEITGKNSNDYLYPEQLKMT 180
DB 276 IALLKIRSGRCAQPSRTIQTICLPSMYNDPOFGTSCEITGKNSNDYLYPEQLKMT 335
QY 181 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDCQDGGPLVCSLQGRMTLTGIVS 240
DB 336 VVKLISHRECQPHYGVSEVTTKMLCAADPQWKTDCQDGGPLVCSLQGRMTLTGIVS 395
QY 241 WGRGCALKDKPGVTVRVSHFLPWRSHRKEENGLAL 276
DB 396 WGRGCALKDKPGVTVRVSHFLPWRSHRKEENGLAL 431

```

```

RESULT 20
US-09-880-503-6
; Sequence 6, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B

```

QY 69 HCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRS 128  
DB 61 HCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRS 120  
QY 129 KEGRCQAQPSRTIQTCLPSMNDPQGTSCITGFKENSTDYLYPEQLKMTVVKLSHR 188  
DB 121 KEGRCQAQPSRTIQTCLPSMNDPQGTSCITGFKENSTDYLYPEQLKMTVVKLSHR 180  
QY 189 EQCQPHYGSEVTTKMLCAADPQWKTSCQDSDGGLVCSLQGRMTLTGIVSWGRGCALK 248  
DB 181 EQCQPHYGSEVTTKMLCAADPQWKTSCQDSDGGLVCSLQGRMTLTGIVSWGRGCALK 240  
QY 249 DRPGVYTRVSHFLPWIRSHKENGIAL 276  
DB 241 DRPGVYTRVSHFLPWIRSHKENGIAL 268

RESULT 22  
US-10-360-101-266  
; Sequence 266, Application US/10360101  
; Publication No. US20040009550A1  
; GENERAL INFORMATION:  
; APPLICANT: Moll, N.  
; APPLICANT: Leenhouts, Cornelis J.  
; TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way  
; FILE REFERENCE: 2183-5673  
; CURRENT APPLICATION NUMBER: US/10/360,101  
; PRIOR FILING DATE: 2003-02-07  
; PRIOR APPLICATION NUMBER: EP 02077060.8  
; PRIOR FILING DATE: 2002-05-24  
; NUMBER OF SEQ ID NOS: 309  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 266  
; LENGTH: 445  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: sequence of urokinase  
US-10-360-101-266

Query Match 96.4%; Score 1453; DB 15; Length 445;  
Best Local Similarity 96.5%; Pred. No. 4.8e-140;  
Matches 276; Conservative 0; Mismatches 0; Indels 10; Gaps 5;

QY 1 KPSPPPEE--LKQCCQKTLRPFKLIIGGEFTTIENQPFPAIYRRHRGGSVTVVCGGSL 58  
DB 160 KPSPPPEEFLKQCCQKTLRPFKLIIGGEFTTIENQPFPAIYRRHRGGSVTVVCGGSL 219  
QY 59 ISPCWVISA--THCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLA 116  
DB 220 ISPCWVISAFTTRCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLA 279  
QY 117 HNDIALKTI--RSKEGRCAQPSRTIOTICLPSMNDPQGTSCITGFKENSTDYLYP 174  
DB 280 HNDIALKTI--RSKEGRCAQPSRTIOTICLPSMNDPQGTSCITGFKENSTDYLYP 339  
QY 175 EQLKMTVVKLI--SHRECOQPHYGSEVTTKMLCAADPQWKTSCQDSDGGLVCSLQGR 232  
DB 340 EQLKMTVVKLI--SHRECOQPHYGSEVTTKMLCAADPQWKTSCQDSDGGLVCSLQGR 399  
QY 233 MTLTGIVSWGRG--CALKDKPGVYTRVSHFLPWIRSHKENGIAL 276  
DB 400 MTLTGIVSWGRGFTALKDKPGVYTRVSHFLPWIRSHKENGIAL 445

RESULT 23  
US-09-264-468B-2  
; Sequence 2, Application US/09264468B  
; Patent No. US20020106775A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Jieyi  
; APPLICANT: Nienaber, Vicki L.  
; APPLICANT: Henkin, Jack

APPLICANT: Smith, Richard A.  
APPLICANT: Walter, Karl A.  
APPLICANT: Severin, Jean M.  
APPLICANT: Edalji, Robinton  
APPLICANT: Johnson Jr., Robert W.  
APPLICANT: Holzman, Thomas F.  
TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE  
FILE REFERENCE: 6310.US.P1  
CURRENT APPLICATION NUMBER: US/09/264,468B  
CURRENT FILING DATE: 1999-03-05  
PRIOR APPLICATION NUMBER: US 09/036,361  
PRIOR FILING DATE: 1998-03-06  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2  
LENGTH: 246  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-264-468B-2

Query Match 88.4%; Score 1333; DB 9; Length 246;  
Best Local Similarity 99.2%; Pred. No. 4.3e-128;  
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0.

QY 24 IIGGEFTTIENQPFPAIYRRHRGGSVTVVCGGSLISPCWVISAHCFIDYPKEDYIVY 83  
DB 1 IIGGEFTTIENQPFPAIYRRHRGGSVTVVCGGSLISPCWVISAHCFIDYPKEDYIVY 60  
QY 84 LGRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSKEGRCAQPSRTIOTI 143  
DB 61 LGRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSKEGRCAQPSRTIOTI 120  
QY 144 CLPSMNDPQGTSCITGFKENSTDYLYPEQLKMTVVKLSHRECOQPHYGSEVTTK 203  
DB 121 ALPSMNDPQGTSCITGFKENSTDYLYPEQLKMTVVKLSHRECOQPHYGSEVTTK 180  
QY 204 MLCAADPQWKTSCQDSDGGLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPW 263  
DB 181 MLCAADPQWKTSCQDSDGGLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPW 240  
QY 264 IRSHTK 269  
DB 241 IRSHTK 246

RESULT 24  
US-09-898-837A-47  
; Sequence 47, Application US/09898837A  
; Publication No. US20030077697A1  
; GENERAL INFORMATION:  
; APPLICANT: Quinn, Kerry E.  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Majumder, Rumud  
; APPLICANT: Vernet, Corine  
; APPLICANT: Herrmann, John L.  
; APPLICANT: Burgess, Catherine  
; APPLICANT: Fernandes, Elma  
; APPLICANT: Taupier Jr., Raymond  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Curagen Corporation  
; APPLICANT: Gerlach, Valerie L.  
; APPLICANT: MacDougall, John R.  
TITLE OF INVENTION: NOVEL SERINE/THREONINE PROTEIN-KINASE LIKE PROTEINS AND  
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING THE SAME  
FILE REFERENCE: 15966-598 CIP  
CURRENT APPLICATION NUMBER: US/09/898,837A  
CURRENT FILING DATE: 2001-07-03  
PRIOR APPLICATION NUMBER: U.S.N. 60/165,986  
PRIOR FILING DATE: 1999-11-17  
PRIOR APPLICATION NUMBER: U.S.N. 60/194,839  
PRIOR FILING DATE: 2000-04-05  
PRIOR APPLICATION NUMBER: U.S.N. 60/195,637  
PRIOR FILING DATE: 2000-04-07

Query Match 87.4%; Score 1318; DB 10; Length 241;  
Best Local Similarity 99.6%; Pred. No. 1.5e-126;  
Matches 240; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
US-09-898-837A-47  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/197,080  
; PRIOR FILING DATE: 2000-04-13  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/232,677  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/181,347  
; PRIOR FILING DATE: 2000-02-09  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/194,195  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/215,906  
; PRIOR FILING DATE: 2000-07-03  
; PRIOR APPLICATION NUMBER: U.S.S.N. 09/715,427  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 47  
; LENGTH: 241  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-898-837A-47  
Query Match 87.4%; Score 1318; DB 10; Length 241;  
Best Local Similarity 99.6%; Pred. No. 1.5e-126;  
Matches 240; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
US-09-898-837A-47  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/197,080  
; PRIOR FILING DATE: 2000-04-13  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/232,677  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/181,347  
; PRIOR FILING DATE: 2000-02-09  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/194,195  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/215,906  
; PRIOR FILING DATE: 2000-07-03  
; PRIOR APPLICATION NUMBER: U.S.S.N. 09/715,427  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 47  
; LENGTH: 241  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-898-837A-47

Query Match 87.4%; Score 1318; DB 10; Length 241;  
Best Local Similarity 99.6%; Pred. No. 1.5e-126;  
Matches 240; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
US-09-898-837A-47  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/197,080  
; PRIOR FILING DATE: 2000-04-13  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/232,677  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/181,347  
; PRIOR FILING DATE: 2000-02-09  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/194,195  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/215,906  
; PRIOR FILING DATE: 2000-07-03  
; PRIOR APPLICATION NUMBER: U.S.S.N. 09/715,427  
; PRIOR FILING DATE: 2000-11-16  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 47  
; LENGTH: 241  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-898-837A-47

Query Match 73.8%; Score 1112.5; DB 12; Length 433;  
Best Local Similarity 71.0%; Pred. No. 3.7e-105;  
Matches 196; Conservative 39; Mismatches 40; Indels 1; Gaps 1;  
US-10-087-192-591

QY 1 KPSSPPBELKFCGQOKTLRPRFKLIIGGFTTIENQPFWFAAIYRRHGGG-VTVYCGGSLI 59  
DB 157 KPSSSVDDQGGFCGQOKALRPRFKIIVGGFTTEVENQPFWFAAIYRRHGGGPPSFKCGGSLI 216  
QY 60 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVEVENILHKDYADTLAHNN 119  
DB 217 PCWVASAAHCFIQPKENYVYLGQSKESYNFGEMKFEVEQLILHEYYREDSLAYHN 276  
QY 120 DIALLKIRSKRGCRQPSRTIOTICLPWYNDPQGTSCETITGFGKENSTDYLYPEQLK 179  
DB 277 DIALLKIRTSQCAQPSRSIQICLPFRFTDAPFGSDCEITGFGKESSEDYLYPRNKK 336  
QY 180 TVVKLISHRECOQPHYVYGVSEVTKMLCAADPQWKTDSQDGGSGPLVCSLQGRMTLTGIV 239  
DB 337 SVVKLVSHQCMQPHYVYGVSEINVKMLCAADPEWKTDSCKDGGGPLICNIEGRPTLSGIV 396  
QY 240 SWGRGCAKDKPGVYTRVSHFLPWIRSHYKENG 275  
DB 397 SWGRGCAENKPGVYTRVSHFLDWIQSHIGERKGLA 432

RESULT 26  
US-10-106-698-6266  
; Sequence 6266, Application US/10106698  
; Publication No. US20030109690A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Pepti  
; FILE REFERENCE: PA005P1  
; CURRENT APPLICATION NUMBER: US/10/106,698  
; PRIOR FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: PCT/US00/26524  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US 60/157,137  
; PRIOR FILING DATE: 1999-09-29  
; PRIOR APPLICATION NUMBER: US 60/163,280  
; PRIOR FILING DATE: 1999-11-03  
; NUMBER OF SEQ ID NOS: 8564  
; SOFTWARE: Patent in Ver. 3.0  
; SEQ ID NO 6266  
; LENGTH: 337  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-106-698-6266  
Query Match 49.3%; Score 742; DB 14; Length 337;  
Best Local Similarity 90.9%; Pred. No. 2.5e-67;  
Matches 140; Conservative 3; Mismatches 5; Indels 6; Gaps 1;  
US-10-106-698-6266

QY 1 KPSSPPBELKFCGQOKTLRPRFKLIIGGFTTIENQPFWFAAIYRRHGGG-VTVYCGGSLI 60  
DB 162 KPSSPPBELKFCGQOKTLRPRFKLIIGGFTTIENQPFWFAAIYRRHGGG-VTVYCGGSLI 221  
QY 61 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVEVENILHKDYADTLAHNN 120  
DB 222 PCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVEVENILHKDYADTLAHNN 281  
QY 121 IALLKIRSKRGCRQPSRTIOTICLPWYNDPQGTSCETITGFGKENSTDYLYPEQLK 148  
DB 282 IALLKIRSKRGCRQPSRTIOTICLPWYNDPQGTSCETITGFGKENSTDYLYPEQLK 315

RESULT 27  
US-10-264-049-2927  
; Sequence 2927, Application US/10264049  
; Publication No. US20040005579A1  
; GENERAL INFORMATION:  
; APPLICANT: Birse et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA133P1  
; CURRENT APPLICATION NUMBER: US/10/264,049  
; PRIOR FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/18569

```
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: Patent In Ver. 3.1
; SEQ ID NO 2927
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-049-2927

Query Match          49.2%; Score 742; DB 15; Length 337;
Best Local Similarity 90.9%; Pred. No. 2.5e-67;
Matches 140; Conservative 3; Mismatches 5; Indels 6; Gaps 1;

QY 1 KPSSPEELKQCGQKTLRPRFKIIGGFTTIENQPFPAALYRRHGGSVTVVCGSLIS 60
DB 162 KPSSPEELKQCGQKTLRPRFKIIGGFTTIENQPFPAALYRRHGGSVTVVCGSLIS 221
QY 61 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 120
DB 222 PCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHND 281
QY 121 IALLKIRSKGRCAQ-----PSRTIQTICLPSM 148
DB 282 IALLKIRSKGRCAQHPGLYRPSACPRCITIPSL 315

RESULT 28
US-09-997-003-39
; Sequence 39, Application US/09997003
; Publication No. US2003020361A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA003P1
; CURRENT APPLICATION NUMBER: US/09/997,003
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: unassigned
; PRIOR FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: PCT/US00/22157
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/148,680
; PRIOR FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 39
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-997-003-39

Query Match          39.2%; Score 591; DB 11; Length 118;
Best Local Similarity 100.0%; Pred. No. 1.9e-52;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 DYLYPEQLKMTVVKLISHRECQPHYVGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSL 229
DB 12 DYLYPEQLKMTVVKLISHRECQPHYVGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSL 71
QY 230 QGRMTLTGIVSWRGKALKDKPGVYTRVSHFLPWIRSHTKENGLAL 276
DB 72 QGRMTLTGIVSWRGKALKDKPGVYTRVSHFLPWIRSHTKENGLAL 118

RESULT 29
US-09-987-457-17
; Sequence 17, Application US/09987457
; Publication No. US2003013150A1
; GENERAL INFORMATION:
; APPLICANT: Manosroi, Aranya
; APPLICANT: Manosroi, Jiradej
; APPLICANT: Tayapiwatana, Chatchai
```

```
; APPLICANT: Goetz, Friedrich
; APPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryote
; FILE REFERENCE: 0652.2180001
; CURRENT APPLICATION NUMBER: US/09/987,457
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,573
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 00 27 782.2
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 17
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: part of the
; OTHER INFORMATION: recombinant K2S molecule (modified)
US-09-987-457-17

Query Match          38.7%; Score 583; DB 10; Length 268;
Best Local Similarity 44.9%; Pred. No. 3.7e-51;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5.

QY 13 CG-QKTLRPRFKIIGGFTTIENQPFPAALYRRH-RGGSVTVVCGSLISPCWVISATHC 70
DB 5 CGLRQYSQPFRIKGLGFADIASHPQAAIFAKHRRSPGFRFLCGGILISSCWILSAHC 64
QY 71 FIDYPKKEDYIVLGRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSK 130
DB 65 FQERFPFHLLTVILGTYRVVPGSEEQKEVEKYIVHKEFDDDT--YNDIALLQLKSDS 122
QY 131 GRCAQPSRTIQTICLPSMNDPQFGTSCBITGFKENSTDYLYPEQLKMTVVKLISHREC 190
DB 123 SRCAQESSVVRTVCLPFPADLPDWTCELSGVGKEALSPFYSELKEAHRVLPSSRC 182
QY 191 QQPHYVGVSEVTTKMLCAAD-----PQWKTDSCQDGGPLVCSLQGRMTLTGIVSWRG 244
DB 183 TSQHLNRTVTDNMLCAGTTRSGGPOANLHDACQDGGPLVCLNDGRMTLVGIISWGLG 242
QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269
DB 243 CGQKDVPGVYTKVTVLDWIRDNMR 267

RESULT 30
US-09-987-455-18
; Sequence 18, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 18
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: K2S 260-527
```



US-09-987-455-18

Query Match 38.7%; Score 583; DB 10; Length 268;  
Best Local Similarity 44.9%; Pred. No. 3.7e-51;  
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

Qy 13 CG-QKTLRPRFKIIGGFTTIENQPFWFAAIYRRH-RGGSVTVYVCGGSLISPCWVISATHC 70  
Db 5 CGLRQYSQPQFRIKGGLFADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAHC 64  
Qy 71 FIDYFKKEDYTVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHNDIALKIRSK 130  
Db 65 FOERFPFPHLTVILGRTYRVVPGEEQKEVEKIVHKEFDDT--YDNDIALQLKSDS 122  
Qy 131 GRCAQPSRTIQTICLPSMYNDPQGTSCETGKENSVDYLYPEQLKMTVVKLISHREC 190  
Db 123 SRCQSSSVVTVCLPPADLQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSRC 182  
Qy 191 QQPHYGSEVTTKMLCAAD-----PWKT-DSQGDGGPLVCLNDGRMTLVGIISWGLG 244  
Db 183 TSQHLNRTVTDNMLCAGDTRSGGPQANLHDACQDGGPLVCLNDGRMTLVGIISWGLG 242  
Qy 245 CALKDKPGVTVRSHFLPWIRSHTK 269  
Db 243 CGQKDVPGVTVKTYNLDWIRDNR 267

RESULT 31

US-09-987-457-16  
; Sequence 16, Application US/09987457  
; Publication No. US20030013150A1  
; GENERAL INFORMATION:  
; APPLICANT: Manosroi, Aranya  
; APPLICANT: Manosroi, Jiradej  
; APPLICANT: Tavapiwatana, Chatchai  
; APPLICANT: Goetz, Friedrich  
; APPLICANT: Werner, Rolf-Guenther  
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes  
; FILE REFERENCE: 0652.2180001  
; CURRENT APPLICATION NUMBER: US/09/987,457  
; PRIOR FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,573  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 00 27 782.2  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 16  
; LENGTH: 308  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: part of the recombinant K2S molecule (modified)

US-09-987-457-16

Query Match 38.7%; Score 583; DB 10; Length 308;  
Best Local Similarity 44.9%; Pred. No. 4.5e-51;  
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

Qy 13 CG-QKTLRPRFKIIGGFTTIENQPFWFAAIYRRH-RGGSVTVYVCGGSLISPCWVISATHC 70  
Db 45 CGLRQYSQPQFRIKGGLFADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAHC 104  
Qy 71 FIDYFKKEDYTVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHNDIALKIRSK 130  
Db 105 FOERFPFPHLTVILGRTYRVVPGEEQKEVEKIVHKEFDDT--YDNDIALQLKSDS 162  
Qy 131 GRCAQPSRTIQTICLPSMYNDPQGTSCETGKENSVDYLYPEQLKMTVVKLISHREC 190  
Db 163 SRCQSSSVVTVCLPPADLQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSRC 222  
Qy 191 QQPHYGSEVTTKMLCAAD-----PWKT-DSQGDGGPLVCLNDGRMTLVGIISWGLG 244

Db 223 TSQHLNRTVTDNMLCAGDTRSGGPQANLHDACQDGGPLVCLNDGRMTLVGIISWGLG 282  
Qy 245 CALKDKPGVTVRSHFLPWIRSHTK 269  
Db 283 CGQKDVPGVTVKTYNLDWIRDNR 307

RESULT 32

US-09-987-455-17  
; Sequence 17, Application US/09987455  
; Publication No. US20030049729A1  
; GENERAL INFORMATION:  
; APPLICANT: Aranya Manosroi  
; APPLICANT: Jiradej Manosroi  
; APPLICANT: Chatchai Tavapiwatana  
; APPLICANT: Friedrich Goetz  
; APPLICANT: Rolf-Guenther Werner  
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant  
; FILE REFERENCE: 0652.2190001  
; CURRENT APPLICATION NUMBER: US/09/987,455  
; PRIOR FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,574  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 0027779.8  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 17  
; LENGTH: 308  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: K2S 220-527

US-09-987-455-17

Query Match 38.7%; Score 583; DB 10; Length 308;  
Best Local Similarity 44.9%; Pred. No. 4.5e-51;  
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

Qy 13 CG-QKTLRPRFKIIGGFTTIENQPFWFAAIYRRH-RGGSVTVYVCGGSLISPCWVISATHC 70  
Db 45 CGLRQYSQPQFRIKGGLFADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAHC 104  
Qy 71 FIDYFKKEDYTVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHNDIALKIRSK 130  
Db 105 FOERFPFPHLTVILGRTYRVVPGEEQKEVEKIVHKEFDDT--YDNDIALQLKSDS 162  
Qy 131 GRCAQPSRTIQTICLPSMYNDPQGTSCETGKENSVDYLYPEQLKMTVVKLISHREC 190  
Db 163 SRCQSSSVVTVCLPPADLQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSRC 222  
Qy 191 QQPHYGSEVTTKMLCAAD-----PWKT-DSQGDGGPLVCLNDGRMTLVGIISWGLG 244  
Db 223 TSQHLNRTVTDNMLCAGDTRSGGPQANLHDACQDGGPLVCLNDGRMTLVGIISWGLG 282  
Qy 245 CALKDKPGVTVRSHFLPWIRSHTK 269  
Db 283 CGQKDVPGVTVKTYNLDWIRDNR 307

RESULT 33

US-09-987-457-11  
; Sequence 11, Application US/09987457  
; Publication No. US20030013150A1  
; GENERAL INFORMATION:  
; APPLICANT: Manosroi, Aranya  
; APPLICANT: Manosroi, Jiradej  
; APPLICANT: Tavapiwatana, Chatchai  
; APPLICANT: Goetz, Friedrich  
; APPLICANT: Werner, Rolf-Guenther  
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes

[illegible]

```

; CHECKED BY SEQ ID NO: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: part of the
; ; OTHER INFORMATION: recombinant K2S molecule (modified)
US-09-987-457-13

Query Match          38.7%; Score 583; DB 10; Length 335;
Best Local Similarity 44.9%; Pred. No. 5e-51;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5

Qy 13 CG-QXTLRPRFKIIGGETTIENQWPAALVRRH-RGGSVTVYVCGGSLSPCWVISAATHC 70
Db   :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
72 CGLRQVSQPFRIKGGFLFADIASHPWQAAIAFAKHRRSPGRFLPCGGIITSSCWILSAHC 131
Qy 71 FDIYPKEDDIYVLCGRSLNNTGEMKFVEENILHLKDYSDATLHAHNDIALALKIRSE 130
Db 132 PQERPPPHLTVILGRTVRVVPGSEEQFEVEKYIVHKEFDDDT--YDNDIALQLQKDS 189
Qy 131 GRCAQPSRTIQTICLPSMYNDPQGTSCETGFGKENSITDLYPEQIKMTVVKLI SHRC 190
Db 190 SRCADSSVVRVTCVLPADQLPDWTWCELSGYGKHEALSPFYSERLKEAHVELYPSRC 249
Qy 191 QQPHYYSVTTKMLCAAD-----PQWKT-DSOQDSSGGLPVCSLQGRMTLTGIVSWGRC 244
Db 250 TSOHLNRTVTNNMLCAGSTRSGSPQANLHPDACQDGGGLVCLNDRGMTLVGLIISWGLG 309

```

```
QY 245 CALKDKPGVYTRVSHFLPWIRSHYK 269
Db 310 CGQKDPGVYTKVTNYLDWIRDMR 334

RESULT 36
US-09-987-455-14
; Sequence 14, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tavapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; FILE REFERENCE: DNA-Derived tpa or K2S Molecules
; CURRENT APPLICATION NUMBER: US/09/987,455
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: part of the
; OTHER INFORMATION: recombinant K2S molecule (modified)
US-09-987-457-12
Query Match 38.7%; Score 583; DB 10; Length 339;
Best Local Similarity 44.9%; Pred. No. 5.le-51;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-QKTLRPRFKIIGCEFTTQWFAAIVRRH-RGGSVTYVCGSLISPCWVISATHC 70
Db 76 CGLRQYSQPFRIKGLFADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAHC 135
QY 71 FIDYPKKEDIYVILGRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSKS 130
Db 136 FQERPPPHLTVILGRYRVVPGEEQKFEVEKYIVHKEFDDT--YDNDIALQLKSDS 193
QY 131 GRCAQPSRTIOTICLPMSYNDPQGTSCETGFGKENSTDYLYPEOLKMTVVKLISHREC 190
Db 194 SRCQSSSVRVVCLPPADLQLPDWTCELSYGKHEALSPFYSERLKEAHVRLYPSRC 253
QY 191 QQPHYGSEVTTKMLCAAD-----POWKT-DSCQDGGPLVCSLQGRMTLTGIVSWG 244
Db 254 TSQHLNRTVTDNMLCAGDTRSGGQANLHDACQDGGPLVCLNDGRMTLVGIISWGLG 313
QY 245 CALKDKPGVYTRVSHFLPWIRSHYK 269
Db 314 CGQKDPGVYTKVTNYLDWIRDMR 338

RESULT 38
US-09-987-455-13
; Sequence 13, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tavapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; FILE REFERENCE: DNA-Derived tpa or K2S Molecules
; CURRENT APPLICATION NUMBER: US/09/987,455
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: K2S 193-527,
; OTHER INFORMATION: modified
US-09-987-455-13
Query Match 38.7%; Score 583; DB 10; Length 339;
Best Local Similarity 44.9%; Pred. No. 5.le-51;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-QKTLRPRFKIIGCEFTTQWFAAIVRRH-RGGSVTYVCGSLISPCWVISATHC 70
Db 72 CGLRQYSQPFRIKGLFADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAHC 131
QY 71 FIDYPKKEDIYVILGRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSKS 130
Db 132 FQERPPPHLTVILGRYRVVPGEEQKFEVEKYIVHKEFDDT--YDNDIALQLKSDS 189
QY 131 GRCAQPSRTIOTICLPMSYNDPQGTSCETGFGKENSTDYLYPEOLKMTVVKLISHREC 190
Db 190 SRCQSSSVRVVCLPPADLQLPDWTCELSYGKHEALSPFYSERLKEAHVRLYPSRC 249
QY 191 QQPHYGSEVTTKMLCAAD-----POWKT-DSCQDGGPLVCSLQGRMTLTGIVSWG 244
Db 250 TSQHLNRTVTDNMLCAGDTRSGGQANLHDACQDGGPLVCLNDGRMTLVGIISWGLG 309
QY 245 CALKDKPGVYTRVSHFLPWIRSHYK 269
Db 310 CGQKDPGVYTKVTNYLDWIRDMR 334

RESULT 37
US-09-987-457-12
; Sequence 12, Application US/09987457
; Publication No. US20030013150A1
; GENERAL INFORMATION:
; APPLICANT: Manosroi, Aranya
; APPLICANT: Manosroi, Jiradej
; APPLICANT: Tavapiwatana, Chatchai
; APPLICANT: Goetz, Friedrich
; APPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 0652.2180001
; CURRENT APPLICATION NUMBER: US/09/987,457
```



```

; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 15
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: K2S 191-527,
; OTHER INFORMATION: modified
US-09-987-455-15

Query Match      38.7%; Score 583; DB 10; Length 343;
Best Local Similarity 44.9%; Pred. No. 5.1e-51;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-QKTLRPRFKIIGGFTTIENQPFQFAAIYRRH-RGGSVTVYCGGSLISPCWVISATHC 70
DB 80 CGLRQYSQPFRIKGGFLFADIAHPQAAIFAKHRRSPGERFLCGGILISSCWILSAHC 139
QY 71 FIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSK 130
DB 140 FQERPPPHLTVILGRYRVVPGEEQFEVEKYIVHKEFDDDT--YNDIALQLKSDS 197
QY 131 GRCAQPSRTIOTICLPMSYNDPQFGTSCBITGFGKENSTDYLYPEQLKMTVVKLISHREC 190
DB 198 SRCAQESSVVRVTVCLPPADLQLPDWTCELSGKGHEALSPFYSERLKEAHVRLYPSRRC 257
QY 191 QQPHYGSEVTTKMLCAAD-----PQWKT-DSQCQDSGGPLVCSLQGRMTLTGIVSWGRG 244
DB 258 TSCHLLNRTVTDNMLCAGTRSGGPOANLHDACQDSGGPLVCLNDGRMTLVGIISWGLG 317
QY 245 CALKDKPGVYTVRSHFLPWIRSHTK 269
DB 318 CGQKDVPGVYTVKTVNLDWIRDNR 342

RESULT 43
US-09-987-457-10
; Sequence 10, Application US/09987457
; Publication No. US20030013150A1
; GENERAL INFORMATION:
; APPLICANT: Manosroi, Aranya
; APPLICANT: Manosroi, Jiradej
; APPLICANT: Tayapiwatana, Chatchai
; APPLICANT: Goetz, Friedrich
; APPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryot
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,457
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,573
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 00 27 782.2
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: part of the
; OTHER INFORMATION: recombinant K2S molecule
US-09-987-457-10

Query Match      38.7%; Score 583; DB 10; Length 354;
Best Local Similarity 44.9%; Pred. No. 5.4e-51;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-QKTLRPRFKIIGGFTTIENQPFQFAAIYRRH-RGGSVTVYCGGSLISPCWVISATHC 70
DB 91 CGLRQYSQPFRIKGGFLFADIAHPQAAIFAKHRRSPGERFLCGGILISSCWILSAHC 150
QY 71 FIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSK 130
DB 151 FQERPPPHLTVILGRYRVVPGEEQFEVEKYIVHKEFDDDT--YNDIALQLKSDS 208
QY 131 GRCAQPSRTIOTICLPMSYNDPQFGTSCBITGFGKENSTDYLYPEQLKMTVVKLISHREC 190
DB 209 SRCAQESSVVRVTVCLPPADLQLPDWTCELSGKGHEALSPFYSERLKEAHVRLYPSRRC 268
QY 191 QQPHYGSEVTTKMLCAAD-----PQWKT-DSQCQDSGGPLVCSLQGRMTLTGIVSWGRG 244
DB 269 TSCHLLNRTVTDNMLCAGTRSGGPOANLHDACQDSGGPLVCLNDGRMTLVGIISWGLG 328
QY 245 CALKDKPGVYTVRSHFLPWIRSHTK 269
DB 318 CGQKDVPGVYTVKTVNLDWIRDNR 342

; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 16
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: K2S 191-527,
; OTHER INFORMATION: modified
US-09-987-455-16

Query Match      38.7%; Score 583; DB 10; Length 343;
Best Local Similarity 44.9%; Pred. No. 5.1e-51;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;

QY 13 CG-QKTLRPRFKIIGGFTTIENQPFQFAAIYRRH-RGGSVTVYCGGSLISPCWVISATHC 70
DB 80 CGLRQYSQPFRIKGGFLFADIAHPQAAIFAKHRRSPGERFLCGGILISSCWILSAHC 139
QY 71 FIDYPKEDYIVYLGSRSLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSK 130
DB 140 FQERPPPHLTVILGRYRVVPGEEQFEVEKYIVHKEFDDDT--YNDIALQLKSDS 197
QY 131 GRCAQPSRTIOTICLPMSYNDPQFGTSCBITGFGKENSTDYLYPEQLKMTVVKLISHREC 190
DB 198 SRCAQESSVVRVTVCLPPADLQLPDWTCELSGKGHEALSPFYSERLKEAHVRLYPSRRC 257
QY 191 QQPHYGSEVTTKMLCAAD-----PQWKT-DSQCQDSGGPLVCSLQGRMTLTGIVSWGRG 244
DB 258 TSCHLLNRTVTDNMLCAGTRSGGPOANLHDACQDSGGPLVCLNDGRMTLVGIISWGLG 317
QY 245 CALKDKPGVYTVRSHFLPWIRSHTK 269
DB 318 CGQKDVPGVYTVKTVNLDWIRDNR 342

RESULT 42
US-09-987-455-16
; Sequence 16, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 16
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: K2S 191-527,
; OTHER INFORMATION: modified
US-09-987-455-16
```

```
Db 329 CGQKDVPGVYTKVTNYLDWIRDNR 353
US-09-880-503-5
RESULT 44
US-09-880-503-5
Sequence 11, Application US/0987455
Publication No. US20030049729A1
GENERAL INFORMATION:
APPLICANT: Aranya Manosroi
APPLICANT: Jiradej Manosroi
APPLICANT: Chatchai Tayapiwatana
APPLICANT: Friedrich Goetz
APPLICANT: Rolf-Guenther Werner
TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
FILE REFERENCE: 0652.2190001
CURRENT APPLICATION NUMBER: US/09/887,455
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/268,574
PRIOR FILING DATE: 2001-02-15
PRIOR APPLICATION NUMBER: GB 0027779.8
PRIOR FILING DATE: 2000-11-14
NUMBER OF SEQ ID NOS: 25
SOFTWARE: Patent in Ver. 2.1
SEQ ID NO 11
LENGTH: 354
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: K2S 174-527
US-09-880-503-5
Query Match 38.7%; Score 583; DB 10; Length 354;
Best Local Similarity 44.9%; Pred. No. 5.4e-51;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;
QY 13 CG-QKTLRPRFKLIIGGEFTTIENQWFAIYRRH-RGGSVTVVCGSLISPCWVISATHC 70
Db 91 CGLRQYSQPFRIKGGFLADIAHPNQAAIFAKHRRSPGFLCGGILISSCWILSAHC 150
QY 71 FIDYPKKEDYIVVYGRSLNSNTQGMKFEVENLIHKDYSADTLAHNDIALKIRSK 130
Db 151 FQERFPFHLTVLGRYTVVPGEEQKFEVEKYIVHKEFDDT--YNDIALQLKSDS 208
QY 131 GRCAQSRITQITCLPSMNDPQFGSCITGFKENSTDYLPQLKMTVVKLISHREC 190
Db 209 SRCAQSSVVRTVCLPDLQPLDPTWTECELSGKGRHEALSPFYSELKEAHVRLYPSRC 268
QY 191 QOPHYGSEVTTKMLCAAD-----POWKT-DSQGDGGPLVCSLQGRMTLTGIVSWGRG 244
Db 269 TSQHLLNRTVTDNMLCAGDTRSGGPOANLHDAQCGSDGGPLVCLNDGRMTLVGIISWGLG 328
QY 245 CALKDKPGVTVRSHLPWIRSHTK 269
Db 329 CGQKDVPGVYTKVTNYLDWIRDNR 353
RESULT 45
US-09-880-491A-3
Sequence 3, Application US/0908449A
Patent No. US20020061576A1
GENERAL INFORMATION:
APPLICANT: MOORE, PAUL A.
APPLICANT: RUBEN, STEVEN M.
APPLICANT: EBNER, REINHARD
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR-LIKE PROTEASE
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: USA
US-09-880-491A-3
Sequence 3, Application US/10102704
Publication No. US20020164768A1
GENERAL INFORMATION:
APPLICANT: Moore et al.
TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protein
FILE REFERENCE: PF378C1
CURRENT APPLICATION NUMBER: US/10/102,704
CURRENT FILING DATE: 2002-03-22
PRIOR APPLICATION NUMBER: 09/084,491
PRIOR FILING DATE: 1998-05-27
PRIOR APPLICATION NUMBER: 60/048,000
PRIOR FILING DATE: 1997-05-28
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patent in version 3.1
SEQ ID NO 3
LENGTH: 372
TYPE: PRT
ORGANISM: Homo sapiens
US-10-102-704-3
Query Match 38.7%; Score 583; DB 9; Length 372;
Best Local Similarity 44.9%; Pred. No. 5.7e-51;
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;
QY 13 CG-QKTLRPRFKLIIGGEFTTIENQWFAIYRRH-RGGSVTVVCGSLISPCWVISATHC 70
Db 109 CGLRQYSQPFRIKGGFLADIAHPNQAAIFAKHRRSPGFLCGGILISSCWILSAHC 168
QY 71 FIDYPKKEDYIVVYGRSLNSNTQGMKFEVENLIHKDYSADTLAHNDIALKIRSK 130
Db 169 FQERFPFHLTVLGRYTVVPGEEQKFEVEKYIVHKEFDDT--YNDIALQLKSDS 226
QY 131 GRCAQSRITQITCLPSMNDPQFGSCITGFKENSTDYLPQLKMTVVKLISHREC 190
Db 227 SRCAQSSVVRTVCLPDLQPLDPTWTECELSGKGRHEALSPFYSELKEAHVRLYPSRC 286
QY 191 QOPHYGSEVTTKMLCAAD-----POWKT-DSQGDGGPLVCSLQGRMTLTGIVSWGRG 244
Db 287 TSQHLLNRTVTDNMLCAGDTRSGGPOANLHDAQCGSDGGPLVCLNDGRMTLVGIISWGLG 346
QY 245 CALKDKPGVTVRSHLPWIRSHTK 269
Db 347 CGQKDVPGVYTKVTNYLDWIRDNR 371
RESULT 46
US-10-102-704-3
Sequence 3, Application US/10102704
Publication No. US20020164768A1
GENERAL INFORMATION:
APPLICANT: Moore et al.
TITLE OF INVENTION: Tissue Plasminogen Activator-Like Protein
FILE REFERENCE: PF378C1
CURRENT APPLICATION NUMBER: US/10/102,704
CURRENT FILING DATE: 2002-03-22
PRIOR APPLICATION NUMBER: 09/084,491
PRIOR FILING DATE: 1998-05-27
PRIOR APPLICATION NUMBER: 60/048,000
PRIOR FILING DATE: 1997-05-28
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patent in version 3.1
SEQ ID NO 3
LENGTH: 372
TYPE: PRT
ORGANISM: Homo sapiens
US-10-102-704-3
```

Query Match 38.7%; Score 583; DB 13; Length 372;  
Best Local Similarity 44.9%; Pred. No. 5.7e-51;  
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;  
QY 13 CG-QKTLRPRFKIIGGFTTIENQPFAPAIYRRH-RGGSVTVYVCGSLISPCWVISATHC 70  
DB 109 CGLRQYSQFRIKGLFADIASHPWQAIAFAKHRRSPGERFLCGGILISSCWILSAHC 169  
QY 71 FIDYPKEDYIVYGRSLNSNTQGMKFEVENILHKDYSADTLAHNDIALKIRSK 130  
DB 169 FQERPPPHLTVILGRYRVVPGEEQKFEVEKIVHKEFDDT--YNDIALQLKSDS 226  
QY 131 GRCAQPSRTIQTICLPFSMYNDPQFGTSCITGFGKNSDYLYPEQLKMTVVKLISHREC 190  
DB 227 SRCAQESSVVRTVCLPPADLQLPDWTECELSGYGKHEALSPFYSERLKEAHRVLPSSRC 286  
QY 191 QOPHYGSEVTTKMLCAAD-----PQWKT-DSQGDGGPLVCSLQGRMTLVGVSWGRG 244  
DB 287 TSQHLNRTVTDMNLCAGTRSGGPOANLHDACQDGGPLVCLNDGRMTLVGVISWGLG 346  
QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269  
DB 347 CGQKDVPGVYTKVTNYLDWIRDNR 371

RESULT 47  
US-09-987-455-8  
; Sequence 8, Application US/09987455  
; Publication No. US20030049729A1  
; GENERAL INFORMATION:  
; APPLICANT: Aranya Manosroi  
; APPLICANT: Jiradej Manosroi  
; APPLICANT: Chatchai Tavapiwatana  
; APPLICANT: Friedrich Goetz  
; APPLICANT: Rolf-Guenther Werner  
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant  
; FILE REFERENCE: 0652.2190001  
; CURRENT APPLICATION NUMBER: US/09/987,455  
; PRIOR APPLICATION NUMBER: 60/268,574  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 8  
; LENGTH: 377  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; OTHER INFORMATION: Description of Artificial Sequence: OmpA-K2S  
; OTHER INFORMATION: fusion protein  
US-09-987-455-8

Query Match 38.7%; Score 583; DB 10; Length 377;  
Best Local Similarity 44.9%; Pred. No. 5.8e-51;  
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;  
QY 13 CG-QKTLRPRFKIIGGFTTIENQPFAPAIYRRH-RGGSVTVYVCGSLISPCWVISATHC 70  
DB 113 CGLRQYSQFRIKGLFADIASHPWQAIAFAKHRRSPGERFLCGGILISSCWILSAHC 172  
QY 71 FIDYPKEDYIVYGRSLNSNTQGMKFEVENILHKDYSADTLAHNDIALKIRSK 130  
DB 173 FQERPPPHLTVILGRYRVVPGEEQKFEVEKIVHKEFDDT--YNDIALQLKSDS 230  
QY 131 GRCAQPSRTIQTICLPFSMYNDPQFGTSCITGFGKNSDYLYPEQLKMTVVKLISHREC 190  
DB 231 SRCAQESSVVRTVCLPPADLQLPDWTECELSGYGKHEALSPFYSERLKEAHRVLPSSRC 290  
QY 191 QOPHYGSEVTTKMLCAAD-----PQWKT-DSQGDGGPLVCSLQGRMTLVGVSWGRG 244

Db 291 TSQHLNRTVTDMNLCAGTRSGGPOANLHDACQDGGPLVCLNDGRMTLVGVISWGLG 350  
QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269  
DB 351 CGQKDVPGVYTKVTNYLDWIRDNR 375  
RESULT 48  
US-09-987-457-18  
; Sequence 18, Application US/09987457  
; Publication No. US20030013150A1  
; GENERAL INFORMATION:  
; APPLICANT: Manosroi, Aranya  
; APPLICANT: Manosroi, Jiradej  
; APPLICANT: Tavapiwatana, Chatchai  
; APPLICANT: Goetz, Friedrich  
; APPLICANT: Werner, Rolf-Guenther  
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryot  
; FILE REFERENCE: 0652.2180001  
; CURRENT APPLICATION NUMBER: US/09/987,457  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/268,573  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: GB 00 27 782.2  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 527  
; TYPE: PRT  
; ORGANISM: Homo sapiens (tPA)  
US-09-987-457-18

Query Match 38.7%; Score 583; DB 10; Length 527;  
Best Local Similarity 44.9%; Pred. No. 9.1e-51;  
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5  
QY 13 CG-QKTLRPRFKIIGGFTTIENQPFAPAIYRRH-RGGSVTVYVCGSLISPCWVISATHC 70  
DB 264 CGLRQYSQFRIKGLFADIASHPWQAIAFAKHRRSPGERFLCGGILISSCWILSAHC 323  
QY 71 FIDYPKEDYIVYGRSLNSNTQGMKFEVENILHKDYSADTLAHNDIALKIRSK 130  
DB 324 FQERPPPHLTVILGRYRVVPGEEQKFEVEKIVHKEFDDT--YNDIALQLKSDS 381  
QY 131 GRCAQPSRTIQTICLPFSMYNDPQFGTSCITGFGKNSDYLYPEQLKMTVVKLISHREC 190  
DB 382 SRCAQESSVVRTVCLPPADLQLPDWTECELSGYGKHEALSPFYSERLKEAHRVLPSSRC 441  
QY 191 QOPHYGSEVTTKMLCAAD-----PQWKT-DSQGDGGPLVCSLQGRMTLVGVSWGRG 244  
DB 442 TSQHLNRTVTDMNLCAGTRSGGPOANLHDACQDGGPLVCLNDGRMTLVGVISWGLG 501  
QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269  
DB 502 CGQKDVPGVYTKVTNYLDWIRDNR 526

RESULT 49  
US-09-987-455-19  
; Sequence 19, Application US/09987455  
; Publication No. US20030049729A1  
; GENERAL INFORMATION:  
; APPLICANT: Aranya Manosroi  
; APPLICANT: Jiradej Manosroi  
; APPLICANT: Chatchai Tavapiwatana  
; APPLICANT: Friedrich Goetz  
; APPLICANT: Rolf-Guenther Werner  
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant  
; FILE REFERENCE: 0652.2190001  
; CURRENT APPLICATION NUMBER: US/09/987,455  
; CURRENT FILING DATE: 2001-11-14

131 ORCAQPSRTIOTICLPSMYNDPQFGTSCETGFGKENSSTLYPEQLKMTVVKLIHREC 190  
132 SRCAQESSVVRVCLPPADLQLPDWTCELSGKGHEALSPFYSERLKEAHVRLYPSRC 441  
191 QCPHYGSEVTTKMLCAAD-----POWKT-DSCQDGGPLVCSLQGRMTLTGIVSWGRG 244  
442 TSQHLLNRTVTDNMLCAGTSGGPOANLHDACQDGGPLVCLNDGRMTLVGIISWGLG 501  
245 CALKDKPGVYTRVSHFLPWIRSHTK 269  
502 CQKDVPGVYTKVNYLDWIRDNR 526

Search completed: May 25, 2004, 15:03:44  
Job time : 53.3138 secs

Query Match 38.7%; Score 583; DB 10; Length 527;  
Best Local Similarity 44.9%; Pred. No. 9.1e-51;  
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;  
QY 13 CG-QKTLRPFKTIIGGFTTIENQPFAPAIYRRH-RGGSVTVCGGSLISPCWVISATHC 70  
DB 264 CGLRQVSPQFRIKGGFLFADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAHC 323  
QY 71 FIDYPKEDYIVYGRSLNNTQGMKEFVENILHKDYSADTLAHNDIALKIRSK 130  
DB 324 FQERFPPHLLTVILGRTRYRVVPGEBEOKFEVYVHKFEFDDT--YDNDIALQLKSDS 381  
QY 131 GRCAQPSRTIOTICLPSMYNDPQFGTSCETGFGKENSSTLYPEQLKMTVVKLIHREC 190  
DB 382 SRCAQESSVVRVCLPPADLQLPDWTCELSGKGHEALSPFYSERLKEAHVRLYPSRC 441  
QY 191 QCPHYGSEVTTKMLCAAD-----POWKT-DSCQDGGPLVCSLQGRMTLTGIVSWGRG 244  
DB 442 TSQHLLNRTVTDNMLCAGTSGGPOANLHDACQDGGPLVCLNDGRMTLVGIISWGLG 501  
QY 245 CALKDKPGVYTRVSHFLPWIRSHTK 269  
DB 502 CQKDVPGVYTKVNYLDWIRDNR 526

RESULT 50  
US-10-432-842-1  
; Sequence 1, Application US/10432842  
; Publication No. US20040071707A1  
; GENERAL INFORMATION:  
; APPLICANT: Veronica A. CARROLL  
; APPLICANT: Adrian L. HARRIS  
; APPLICANT: Roy BICKNELL  
; APPLICANT: Pat PRICE  
; TITLE OF INVENTION: MODULATION OF CELL GROWTH  
; FILE REFERENCE: 117-450 / N.79507A SER  
; CURRENT APPLICATION NUMBER: US/10/432,842  
; CURRENT FILING DATE: 2003-09-27  
; PRIOR APPLICATION NUMBER: PCT/GB01/05244  
; PRIOR FILING DATE: 2001-11-28  
; PRIOR APPLICATION NUMBER: GB 0029001.5  
; PRIOR FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: MS Word  
; SEQ ID NO 1  
; LENGTH: 527  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-432-842-1

Query Match 38.7%; Score 583; DB 12; Length 527;  
Best Local Similarity 44.9%; Pred. No. 9.1e-51;  
Matches 119; Conservative 39; Mismatches 97; Indels 10; Gaps 5;  
QY 13 CG-QKTLRPFKTIIGGFTTIENQPFAPAIYRRH-RGGSVTVCGGSLISPCWVISATHC 70  
DB 264 CGLRQVSPQFRIKGGFLFADIASHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAHC 323  
QY 71 FIDYPKEDYIVYGRSLNNTQGMKEFVENILHKDYSADTLAHNDIALKIRSK 130  
DB 324 FQERFPPHLLTVILGRTRYRVVPGEBEOKFEVYVHKFEFDDT--YDNDIALQLKSDS 381



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2004, 14:48:05 ; Search time 19.1096 Seconds  
(without alignments)  
745.636 Million cell updates/sec

Title: US-09-880-503-5  
Sequence: 1 KPSSPPPELKFQCCQKTLRP.....VSHFLPWIRSHKTEENGLAL 276

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : Issued Patents AA.\*

1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1508	100.0	365	1	US-08-093-741-83
2	1508	100.0	365	1	US-08-720-012-83
3	1508	100.0	393	2	US-08-560-098A-44
4	1508	100.0	393	3	US-08-967-024C-24
5	1508	100.0	393	3	US-08-967-024C-25
6	1508	100.0	411	1	US-08-087-163-1
7	1508	100.0	411	1	US-08-286-748B-18
8	1508	100.0	411	1	US-08-153-799-18
9	1508	100.0	411	3	US-09-181-816-1
10	1508	100.0	430	6	5219569-2
11	1508	100.0	431	4	US-09-101-272G-1
12	1508	100.0	431	6	5188829-1
13	1508	100.0	432	2	US-08-560-098A-47
14	1505	99.8	411	4	US-09-403-736-2
15	1505	99.8	430	1	US-07-942-157A-3
16	1496	99.2	306	2	US-08-560-098A-45
17	1496	99.2	331	2	US-08-560-098A-46
18	1495	99.1	411	2	US-08-560-098A-48
19	1382	91.6	253	3	US-08-944-483-73
20	1374	91.1	254	2	US-08-560-098A-49
21	607	40.3	355	2	US-08-811-949-59
22	590	39.1	437	2	US-08-811-949-47
23	590	39.1	437	2	US-08-811-949-51
24	590	39.1	527	2	US-08-811-949-33
25	587	38.3	437	2	US-08-811-949-57
26	583	38.7	355	1	US-08-137-116-1
27	583	38.7	355	1	US-08-217-618-1

28	583	38.7	355	1	US-08-427-640-2
29	583	38.7	355	1	US-08-427-640-6
30	583	38.7	355	1	US-08-217-617A-1
31	583	38.7	355	1	US-08-217-616-1
32	583	38.7	355	2	US-08-811-949-45
33	583	38.7	355	3	US-08-794-528-1
34	583	38.7	355	6	522256-1
35	583	38.7	355	1	US-08-427-640-4
36	583	38.7	356	1	US-08-427-640-8
37	583	38.7	378	4	US-09-553-498-10
38	583	38.7	378	4	US-09-618-869-10
39	583	38.7	383	2	US-08-558-269-6
40	583	38.7	383	3	US-09-410-882-6
41	583	38.7	472	2	US-08-811-949-63
42	583	38.7	527	1	US-07-609-510B-16
43	583	38.7	527	5	PCT-US91-01025A-2
44	583	38.7	527	6	5185259-8
45	583	38.7	562	2	US-08-811-949-43
46	583	38.7	562	2	US-08-560-098A-50
47	583	38.7	562	2	US-08-883-795A-38
48	583	38.7	562	4	US-09-703-695A-4
49	583	38.7	562	6	5185259-3
50	583	38.7	562	6	5200340-2
51	583	38.7	562	6	5244676-5
52	583	38.7	562	6	5344773-2
53	582	38.6	527	6	5520913-1
54	580	38.5	355	2	US-08-811-949-53
55	580	38.5	437	2	US-08-811-949-55
56	579.5	38.4	389	2	US-08-811-949-67
57	579	38.4	437	2	US-08-811-949-49
58	578	38.3	389	2	US-08-811-949-65
59	565.5	37.5	253	2	US-09-047-337-8
60	565.5	37.5	253	4	US-09-644-600-8
61	565.5	37.5	253	4	US-09-654-600A-8
62	565.5	37.5	354	2	US-08-811-949-61
63	564.5	37.4	347	2	US-08-811-949-1
64	563.5	37.4	252	3	US-08-944-483-72
65	545	36.1	477	2	US-08-560-098A-51
66	493.5	32.7	300	1	US-08-148-910-1
67	493.5	32.7	300	1	US-08-148-937A-1
68	493.5	32.7	655	1	US-08-148-910-12
69	493.5	32.7	655	1	US-08-448-537A-12
70	476	31.6	243	3	US-08-944-483-70
71	472.5	31.3	248	3	US-08-944-483-71
72	435.5	28.9	284	4	US-09-387-375-7
73	428.5	28.4	316	4	US-09-387-375-9
74	427	28.3	326	4	US-09-411-977-3
75	423.5	28.1	418	1	US-08-508-448C-25

## ALIGNMENTS

RESULT 1  
US-08-093-741-83  
; Sequence 83, Application US/08093741  
; Patent No. 5681721  
; GENERAL INFORMATION:  
; APPLICANT: STEFFENS, GERT J.  
; APPLICANT: WENDT, STEPHAN  
; APPLICANT: SCHNEIDER, JOHANNES  
; APPLICANT: HEINZEL-WIELAND, REGINA  
; APPLICANT: SAUNDERS, DEREK J.  
; TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN INHIBITING EFFECT  
; TITLE OF INVENTION: INHIBITING EFFECT  
; NUMBER OF SEQUENCES: 83  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Evenson, McKee, Edwards & Lenahan  
; STREET: 1200 G Street, N. W. Suite 700  
; CITY: Washington, D.C.  
; COUNTRY: U.S.  
; ZIP: 20005

```

FT DISULFID 66 75 BY SIMILARITY.
FT DISULFID 83 94 BY SIMILARITY.
FT DISULFID 88 105 BY SIMILARITY.
FT DISULFID 107 116 BY SIMILARITY.
FT DISULFID 124 205 BY SIMILARITY.
FT DISULFID 145 187 BY SIMILARITY.
FT DISULFID 176 200 BY SIMILARITY.
FT DISULFID 213 294 BY SIMILARITY.
FT DISULFID 234 276 BY SIMILARITY.
FT DISULFID 265 289 BY SIMILARITY.
FT DISULFID 297 428 INTERCHAIN (BY SIMILARITY).
FT DISULFID 340 356 BY SIMILARITY.
FT DISULFID 348 417 BY SIMILARITY.
FT DISULFID 442 516 BY SIMILARITY.
FT DISULFID 474 490 BY SIMILARITY.
FT DISULFID 506 534 BY SIMILARITY.
FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 260 260 G -> A (IN REF. 1).
FT CONFLICT 325 325 P -> A (IN REF. 1).
SQ SEQUENCE 559 AA; 63122 MW; 8CCEB2BD94514D9 CRC64;

Query Match.
Best Local Similarity 36.5%; Score 824.5; DB 1; Length 559;
Matches 175; Conservative 63; Mismatches 150; Indels 105; Gaps 9;

Qy 3 ELHOVP---SNCDCLNGGTCVSNKYPSNIHWNCNPKFGGQHCEDKSKTCVGGNGHYF 58
Db 74 QCSVPVRSCEBPCFNGGTCQALYPSDF-VCCPDGFGFKKCDIDTRATCFEEQGIT 132
Qy 59 RGKASTMTGRCLPWNASATVLOQTVYHRSADALQLGKHNYCRNPDRNRRCPCVQVQ 118
Db 133 RGTWSTAESAECINMNSVLSLKPYNARPNALKLGHNHYCRNPDRDLKPCWVFKA 192
Qy 119 LKPLVQSCVHDCADGKLGK----- 137
Db 193 GKYTEFCSTPACPKGKSEDCYVGKGYVYRGTHSLTTSQASCLPWNISIVMGKSYTAWT 252
Qy 138 -----PQCG-QKTLRPFKIIG 154
Db 253 NSQALGLGRHNYCRNPGDARPWCHVMKDKLTWEYCDMSPCSTCGLRQVQRPQFRIKG 312
Qy 155 EFTTIENQPFPAAY-RRHGGSVYVYCGSLSPCWVISATHCFDIYPKKEDIYVVLGR 213
Db 313 LYDTITSHPOAPIFVKNKNSPGRFUCGGVGLISSCWVLSAAHCFLEPFPNHLKVVGLR 372
Qy 214 SRLNSNTQGMKFEVENLILHKYSADTLAHNDIALLKIRSKBGRCAQPSRTIQTICLP 273
Db 373 TYRVVPGEEETPEIKYIVHERBDDT--YNDIALQLRSQSKCAQESSVGTACLP 430
Qy 274 SMYNDPQF-----GTSCEITGFGKENSTDYLPQLKMTVVKLSHRECCQPHYVGEVTT 329
Db 431 ----DPNLQLPDWTCELSGVRGHEASPPFSRDLKEAHVRLYFSRCTSQHLFNKVTYN 486
Qy 330 KMLCAADP-----QWKTDSCQDGGSLVCSLQGRMTLTGIVSWGRGCALKDPGVYVTV 384
Db 487 NMLCAGDTRSGNDLHDACQDGGSLVCSLQGRMTLTGIIISWGLGCGQKDPGVYVTV 546
Qy 385 SHFLPWRSHRKYE 397
Db 547 TNYLDWIHNMKG 559

RESULT 14
TPA_BOVIN STANDARD; PRT; 566 AA.
AC Q28158;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator).
GN PLAT.

```

```

OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Ravn P., Berglund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators uPA
RL and tPA."
RL Int. Dairy J. 5:605-617(1995).
CC -|- FUNCTION: Converts the abundant, but inactive, zymogen plasminogen
CC to plasmin by hydrolyzing a single Arg-Val bond in plasminogen. By
CC controlling plasmin-mediated proteolysis, it plays an important
CC role in tissue remodeling and degradation, in cell migration and
CC many other physiological events.
CC -|- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -|- SUBUNIT: Heterodimer of chain A and chain B held by a disulfide
CC bond.
CC -|- SUBCELLULAR LOCATION: Secreted; extracellular.
CC -|- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC ARG-314 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC -|- MISCELLANEOUS: Binds to the kringle structure of the fibrin A
CC chain. Binding to fibrin enhances its catalytic activity.
CC -|- SIMILARITY: Belongs to peptidase family S1.
CC -|- SIMILARITY: Contains 1 EGF-like domain.
CC -|- SIMILARITY: Contains 1 fibronectin type I domain.
CC -|- SIMILARITY: Contains 2 kringle domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; X85800; CAA59795.1; -.
CC HSP; P00750; IRTF.
CC MEROPS; S01.232; -.
CC InterPro; IPR009003; Cys Ser_trypsin.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR000083; Fibrinctnl.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Peptidase_S1.
CC InterPro; IPR001314; Peptidase_S1A.
CC Pfam; PF00008; EGF; 1.
CC Pfam; PF00039; fnl; 1.
CC Pfam; PF00051; kringle; 2.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 2.
CC SMART; SM00181; EGF; 1.
CC SMART; SM00058; FN1; 1.
CC SMART; SM00130; KR; 2.
CC SMART; SM00020; Tryp_Spc; 1.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS00026; EGF_3; 1.
CC PROSITE; PS01253; FIBRONECTIN_1; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS00070; KRINGLE_2; 2.
CC PROSITE; PS00240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC Plasma; Kringle; EGF-like domain; Serine protease; Glycoprotein;
CC plasminogen activator; Hydrolase; Repeat; Signal.
CC SIGNAL 1 21 BY SIMILARITY.
PT

```

QY 121 PLVQEMVHDCADGKLFQCG-OKTLRPFKLIIGBEFTIENOPFAAIYRRHGG-5T 178  
 DB 199 FTSESCSVPCS-----KATCGLRKYKBPQLHSTGGLDITSHPWQAIPAQNRRSSGER 254  
 QY 179 YVCGSLSPQWVIGATHCFIDYPKEDYI-----VILGRSLNNTQGMKFEVENLIL 233  
 DB 255 FLCGGILISSCVLTAHCF-----QESVLPOLKVLGRVYKVPGBEEQTFKVKYKIV 309  
 QY 234 HKDYSGADTLAHNDIALKIRSGRCAPQSTIQTICLPSMYNDPQFGTSCETGFOKE 293  
 DB 310 HKEFFDDT--YNNDIALLQKSDSPCAQESVRAICLPEANLQLPDWTCELSGYKH 367  
 QY 294 NSTDXYLPOLKMTVVKLISHRECOQPHYGVSEVTKMLCAADPQWKT-----DSCQGD 347  
 DB 368 KSSSPYSQLEKHEGVLVPSRCAPKFLFNKTVNNMLCAGDTSRSGIYPNVHDACQGD 427  
 QY 348 SGGPLVCSLOGRWTLTGIVSWRGCGCALDKPKGVYTRVSHFLPWIRSH 394  
 DB 428 SGGPLVCMNDNHTLGIISWVGCGEKDVPGVYTKVTNYLGWIRDN 474

RESULT 13  
 TPA MOUSE  
 ID TPA MOUSE STANDARD; PRT; 559 AA.  
 AC P11214; Q91VP2;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)  
 DE (t-PA) (t-plasminogen activator).  
 GN PIAT.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88087303; PubMed=2826484;  
 RA Rickles R.J., Darrow A.L., Strickland S.;  
 RT "Molecular cloning of complementary DNA to mouse tissue plasminogen  
 RT activator mRNA and its expression during F9 teratocarcinoma cell  
 RT differentiation.";  
 RL J. Biol. Chem. 263:1563-1569 (1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Mammary gland;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Beasak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Whiting J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Blakesley R.W., Tomchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Groomwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smallos D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 CC -!- FUNCTION: Converts the abundant, but inactive, zymogen plasminogen  
 CC to plasmin by hydrolyzing a single Arg-Val bond in plasminogen. By  
 CC controlling plasmin-mediated proteolysis, it plays an important  
 CC role in tissue remodeling and degradation, in cell migration and  
 CC many other physiological events.  
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-Val bond in

CC plasminogen to form plasmin.  
 CC -!- SUBUNIT: Heterodimer of chain A and chain B held by a disulfide  
 CC bond.  
 CC -!- SUBCELLULAR LOCATION: Secreted; extracellular.  
 CC -!- PM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER  
 CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER  
 CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.  
 CC -!- MISCELLANEOUS: Binds to the kringle structure of the fibrin A  
 CC chain. Binding to fibrin enhances its catalytic activity.  
 CC -!- SIMILARITY: Belongs to peptidase family S1.  
 CC -!- SIMILARITY: Contains 1 EGF-like domain.  
 CC -!- SIMILARITY: Contains 1 fibronectin type I domain.  
 CC -!- SIMILARITY: Contains 2 kringle domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboratio  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on it  
 CC use by non-profit institutions as long as its content is in no wa  
 CC modified and this statement is not removed. Usage by and for commercia  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL; J03520; AAA04070.1; --  
 CC EMBL; BC011256; ARH11256.1; --  
 CC PIR; A29941; A29941.  
 CC HSSP; P00750; IASH.  
 CC MEROPS; S01.232; --  
 CC MGD; MGI:97610; Plat.  
 CC InterPro; IPR003003; Cys Ser trypsin.  
 CC InterPro; IPR008209; EGF\_like.  
 CC InterPro; IPR000083; Fibrinctnl.  
 CC InterPro; IPR006210; IEGF.  
 CC InterPro; IPR000001; Kringle.  
 CC InterPro; IPR001254; Peptidase S1.  
 CC InterPro; IPR001314; Peptidase\_S1A.  
 CC Pfam; PF00003; EGF; 1.  
 CC Pfam; PF00033; fn1; 1.  
 CC Pfam; PF00051; kringle; 2.  
 CC Pfam; PF00089; trypsin; 1.  
 CC PRINTS; PR00722; CHYMOTRYPSIN.  
 CC PRINTS; PR00018; KRINGLE.  
 CC ProDom; PD000395; Kringle; 2.  
 CC SMART; SM00181; EGF; 1.  
 CC SMART; SM00058; FN1; 1.  
 CC SMART; SM00130; KR; 2.  
 CC SMART; SM00020; Tryp\_Spc; 1.  
 CC PROSITE; PS00022; EGF\_1; 1.  
 CC PROSITE; PS01186; EGF\_2; 1.  
 CC PROSITE; PS00026; EGF\_3; 1.  
 CC PROSITE; PS01253; FIBRONECTIN\_1; 1.  
 CC PROSITE; PS00021; KRINGLE\_1; 2.  
 CC PROSITE; PS00070; KRINGLE\_2; 2.  
 CC PROSITE; PS02040; TRYPSIN\_DOM; 1.  
 CC PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 CC PROSITE; PS00135; TRYPSIN\_SER; 1.  
 CC Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
 CC Plasma; Kringle; EGF-like domain; Repeat; Signal.  
 CC SIGNAL 1 17  
 CC PROPEP 18 29  
 CC CHAIN 30 559  
 CC CHAIN 30 308  
 CC CHAIN 309 559  
 CC CHAIN 36 78  
 CC DOMAIN 79 117  
 CC DOMAIN 124 205  
 CC DOMAIN 213 294  
 CC DOMAIN 309 559  
 CC DOMAIN 355 355  
 CC ACT\_SITE 404 404  
 CC ACT\_SITE 510 510  
 CC DISULFID 38 68

RA Renatus M., Stubbs M.T., Huber R., Bringmann P., Donner P.,  
 RA Schleuning W.D., Bode W.,  
 RT "Catalytic domain structure of vampire bat plasminogen activator: a  
 RT molecular paradigm for proteolysis without activation cleavage.",  
 RL Biochemistry 36:13483-13493(1997).  
 CC -1- FUNCTION: Probably essential to support the feeding habits of this  
 CC exclusively haematophagous animal. Potent thrombolytic agent.  
 CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
 CC plasminogen to form plasmin.  
 CC -1- ENZYME REGULATION: Activity toward plasminogen is stimulated in  
 CC the presence of fibrin I.  
 CC -1- SUBUNIT: Monomer.  
 CC -1- DOMAIN: The fibronectin type-I domain mediates binding to fibrin,  
 CC and the kringle domain apparently mediates fibrin-induced  
 CC stimulation of activity.  
 CC -1- SIMILARITY: Belongs to peptidase family S1.  
 CC -1- SIMILARITY: Contains 1 EGF-like domain.  
 CC -1- SIMILARITY: Contains 1 fibronectin type I domain.  
 CC -1- SIMILARITY: Contains 1 kringle domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M63987; AAA31591.1; -;  
 DR EMBL; M63986; AAA31592.1; -;  
 DR PIR; JS0597; JS0597.  
 DR PDB; 1A5I; 23-MAR-99.  
 DR MEROPS; S01-232; -;  
 DR GlycoSuiteDB; P98119; -;  
 DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
 DR InterPro; IPR006209; EGF\_like.  
 DR InterPro; IPR000083; Fibrinctn1.  
 DR InterPro; IPR006210; IEGF.  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR Pfam; PF00008; EGF; 1.  
 DR Pfam; PF00035; fnl; 1.  
 DR Pfam; PF00051; kringle; 1.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PRINTS; PR00018; KRINGLE.  
 DR ProDom; PD000395; Kringle; 1.  
 DR SMART; SM00181; EGF; 1.  
 DR SMART; SM00058; FN1; 1.  
 DR SMART; SM00130; KR; 1.  
 DR SMART; SM00020; Tryp\_SPC; 1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS00026; EGF\_3; 1.  
 DR PROSITE; PS02253; FIBRONECTIN\_1; 1.  
 DR PROSITE; PS00021; KRINGLE\_1; 1.  
 DR PROSITE; PS00070; KRINGLE\_2; 1.  
 DR PROSITE; PS0240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
 KW Kringle; EGF-like domain; Signal; Multigene family; 3D-structure.  
 FT SIGNAL 1 36  
 FT CHAIN 37 477 SALIVARY PLASMINOGEN ACTIVATOR ALPHA 1.  
 FT DOMAIN 40 82 FIBRONECTIN TYPE-I.  
 FT DOMAIN 83 121 EGF-LIKE.  
 FT DOMAIN 128 209 KRINGLE.  
 FT DOMAIN 225 477 SERINE PROTEASE.  
 FT ACT\_SITE 272 272 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 321 321 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 428 428 CHARGE RELAY SYSTEM.  
 FT DISULFID 42 72 BY SIMILARITY.

FT DISULFID 79 79  
 FT DISULFID 87 98  
 FT DISULFID 92 109  
 FT DISULFID 111 120  
 FT DISULFID 128 209  
 FT DISULFID 149 191  
 FT DISULFID 180 204  
 FT DISULFID 214 345  
 FT DISULFID 257 273  
 FT DISULFID 265 334  
 FT DISULFID 359 434  
 FT DISULFID 391 407  
 FT DISULFID 424 452  
 FT CARBOHYD 153 153  
 FT CARBOHYD 398 398  
 FT TURN 214 215  
 FT STRAND 223 224  
 FT TURN 226 227  
 FT STRAND 230 231  
 FT HELIX 234 236  
 FT TURN 238 239  
 FT STRAND 240 245  
 FT STRAND 254 263  
 FT TURN 264 265  
 FT STRAND 266 269  
 FT HELIX 271 273  
 FT TURN 280 282  
 FT STRAND 284 287  
 FT TURN 297 298  
 FT STRAND 300 309  
 FT TURN 311 312  
 FT TURN 315 317  
 FT TURN 319 320  
 FT STRAND 323 328  
 FT STRAND 338 338  
 FT TURN 339 340  
 FT STRAND 341 341  
 FT STRAND 343 345  
 FT TURN 349 350  
 FT TURN 355 356  
 FT STRAND 358 363  
 FT STRAND 366 366  
 FT STRAND 374 374  
 FT STRAND 379 385  
 FT STRAND 388 390  
 FT TURN 393 398  
 FT TURN 403 404  
 FT STRAND 405 409  
 FT TURN 425 426  
 FT TURN 428 429  
 FT STRAND 431 436  
 FT TURN 437 438  
 FT STRAND 439 448  
 FT TURN 455 456  
 FT STRAND 459 463  
 FT HELIX 464 467  
 FT HELIX 468 474  
 FT SEQUENCE 477 AA; 53616 MW; AA06FD1739C10E5E CRC64;

Query Match 37.2%; Score 840.5; DB 1; Length 477;  
 Best Local Similarity 43.0%; Pred. No. 2,8e-60;

Matches 175; Conservative 56; Mismatches 147; Indels 29; Gaps 10;

QY 5 HQVPEN-CD---CLNGGTCVSNKYFSNIHWCNCPKKFGQHCIEDKSKTCYEGNCHFYRG 60

Db 80 HTVPVNSCSEPRCFNGGTCQWAVYFSDF--VCQCPAGYTGKRCVDTRATCYEGQGVYRG 138

QY 61 KASDTDMGEPCLPMSATVLTQOTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120

Db 139 TWSTAESRVECIWNSSLLTRRTYNGRMPDAPNLGLNHNCRNPNGAPKPCYVIKACK 198

```

DR PIR; A35029; A35029.
DR HSP; P00750; IRTF.
DR MEROPS; S01.232; -.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibriactn1.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Peptidase S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00039; fni_1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO00722; CHYMOTRYPSIN.
DR PRINTS; PRO0018; KRINGLE.
DR PROSITE; PD000395; Kringle; 2.
DR SMART; SM00181; EGF_1.
DR SMART; SM00058; FNI; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; ITP_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS00240; TRYPsin DOM; 1.
DR PROSITE; PS00134; TRYPsin HIS; 1.
DR PROSITE; PS00135; TRYPsin SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 17 PROBABLE.
FT PROPEP 18 29
FT CHAIN 30 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 30 308 TISSUE-TYPE PLASMINOGEN ACTIVATOR A
FT CHAIN 309 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR B
FT CHAIN 36 78 FIBRONECTIN TYPE-I.
FT DOMAIN 79 117 EGF-LIKE.
FT DOMAIN 124 205 EGF-LIKE.
FT DOMAIN 213 294 KRINGLE 1.
FT DOMAIN 309 559 KRINGLE 2.
FT ACT_SITE 355 355 SERINE PROTEASE.
FT ACT_SITE 404 404 CHARGE RELAY SYSTEM.
FT ACT_SITE 510 510 CHARGE RELAY SYSTEM.
FT DISULFID 28 68 BY SIMILARITY.
FT DISULFID 66 75 BY SIMILARITY.
FT DISULFID 83 94 BY SIMILARITY.
FT DISULFID 88 105 BY SIMILARITY.
FT DISULFID 107 116 BY SIMILARITY.
FT DISULFID 124 205 BY SIMILARITY.
FT DISULFID 145 187 BY SIMILARITY.
FT DISULFID 176 200 BY SIMILARITY.
FT DISULFID 213 294 BY SIMILARITY.
FT DISULFID 234 276 BY SIMILARITY.
FT DISULFID 265 289 BY SIMILARITY.
FT DISULFID 297 428 INTERCHAIN (BY SIMILARITY).
FT DISULFID 340 356 BY SIMILARITY.
FT DISULFID 348 417 BY SIMILARITY.
FT DISULFID 442 516 BY SIMILARITY.
FT DISULFID 474 490 BY SIMILARITY.
FT DISULFID 506 534 BY SIMILARITY.
FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 380 380 E -> K (IN REF. 1).
SQ SEQUENCE 559 AA; 62903 MW; 7DBD3809C1D1C921 CRC64;
Query Match 37.3%; Score 841.5; DB 1; Length 559;
Best local Similarity 36.5%; Pred. No. 2.8e-60;
Matches 180; Conservative 65; Mismatches 143; Indels 105; Gaps 12;

```

```

QY 3 ELHVP-----SNCDCLNGGTCVSNKYFSNIHWCNPKKFGGQHQCEIDKSKTCYEGNGHFY 58
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 74 QCHSVPRVSCSEPRCFNGGTCQALYFSDP-VQCQPDGFGVKCDIDTRATCFEGGQITY 132
QY 59 RGRASDTTMRPCLPWNASATVLOCTTHAHRSDALQLGKHNYCRPNRRRWCYV-QV 117
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 133 RGTWSTAENGABECINWNSSALSQKPYASRPNAIKLGLGNHNYCRNPDVDKPCWYVKA 192
QY 118 GL-----KPLVQECWV----- 128
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 193 GKYTFEFCSTPACPKGPTEDCYVGKGYRGTHSFTTSKASCLPWNMSMILIGKTYTAWRA 252
QY 129 -----HDCADGKLKFO-----CGQKTLR-PRFKLIIG 154
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 253 NSQALGLGRHNYCRNPDGDAKPKWCKDKLTWEYCDMSPCSTCGLRQYKQKQPFRIKGG 312
QY 155 EFTTIENOPFAALY-RRHRGGSVTVYVCGGSLSPCWVISATCFDYPKEDIYVVLGR 213
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 313 LFTDITSHPWQAALFYVKNKSPGERFLCGGVLISSCWNLSAAHCFVERPPPHLKVILGR 372
QY 214 SRLNSNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSKEGRCQPSRTIQICLP 273
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 373 TYRVVPGEEETFEIEKYIVHKEFDDT--YNDIALQLRSDSSQCAQESSVGTACLP 430
QY 274 SMYNDPQF---GTSCEITGFGKENSTDYLYPQLKMTVVKLISHRECQCPHYVGSVTT 329
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 431 ----DPDVQLPDWTECELSGIGKHEASSPFSRLKEAHVRLYPSRSTCTSHLFNKTIIS 486
QY 330 KMLCAADP-----QWKTDSCQSDSGGGLVCSLGRMTLTGIVSGRGKALCKDPGVYTRV 384
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 487 NMLCAGDTRTGGNQDVHDAQCGDGGGLVCMVKRWTLTGIIISWGLGCGGKDPGIYTKV 546
QY 385 SHFLPWIRSHYKE 397
DQ :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DQ 547 TNYLWTIQDNWKQ 559

RESULT 12
URTL DESRO STANDARD; PRT; 477 AA.
AC P98119,1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Salivary plasminogen activator alpha 1 precursor (EC 3.4.21.68) (DSPA
DE alpha-1).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RC MEDLINE=92039036; PubMed=1937019;
RA Kraetschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schlemming W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]
RP CHARACTERIZATION.
RP MEDLINE=93393059; PubMed=1309059;
RA Schlemming W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RC TISSUE=Salivary gland;
RC MEDLINE=98022741; PubMed=9354616;

```



[2] SEQUENCE FROM N.A.  
 RP TISSUE=Fetal lung;  
 RX MEDLINE=88262579; PubMed=3133640;  
 RA Sasaki H., Saito Y., Hayashi M., Otsuka K., Niwa M.;  
 RT "Nucleotide sequence of the tissue-type plasminogen activator cDNA  
 from human fetal lung cells."; (1988).  
 RL Nucleic Acids Res. 16:5695-5695 (1988).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88054470; PubMed=2824147;  
 RA Reddy V.B., Garramone A.J., Sasak H., Wei C.-M., Watkins P., Galli J.,  
 RA Hsiung N.;  
 RT "Expression of human uterine tissue-type plasminogen activator in  
 mouse cells using BpV vectors."; (1987).  
 RL DNA 6:461-472 (1987).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96196143; PubMed=3009482;  
 RA Friezen Degen S.J., Rajput B., Reich E.;  
 RT "The human tissue plasminogen activator gene";  
 RL J. Biol. Chem. 261:6972-6985 (1986).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=84298137; PubMed=5089198;  
 RA NY T., Elgh F., Lund B.;  
 RT "The structure of the human tissue-type plasminogen activator gene:  
 correlation of intron and exon structures to functional and  
 structural domains."; (1984).  
 RL Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359 (1984).  
 RN [6]  
 RP SEQUENCE FROM N.A. (ISOFORM SHORT).  
 RX MEDLINE=96284200; PubMed=3090401;  
 RA Harris T.J., Patel T., Maxson F.A., Little S., Emtage J.S., P.;  
 RA Opdenakker G., Volckaert G., Ronbouts W., Billiau A., Somer P.;  
 RT "Cloning of cDNA coding for human tissue-type plasminogen activator  
 and its expression in *Escherichia coli*."; (1986).  
 RL Mol. Biol. Med. 3:279-292 (1986).  
 RN [7]  
 RP SEQUENCE FROM N.A. (ISOFORM SHORT).  
 RX MEDLINE=90192129; PubMed=2107528;  
 RA Siebert P.D., Fong K.;  
 RT "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from:  
 human endothelial cells."; (1990).  
 RL Nucleic Acids Res. 18:1086-1086 (1990).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bonfield G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalley D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences."; (2002).  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [9]  
 RP SEQUENCE OF 212-361 FROM N.A.  
 RX MEDLINE=83169656; PubMed=6572897;

RA Edlund T., Ny T., Raanby M., Heden L.-O., Palm G., Holmgren E.,  
 RA Josephson S.;  
 RT "Isolation of cDNA sequences coding for a part of human tissue  
 plasminogen activator."; (1983).  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:349-352 (1983).  
 RN [10]  
 RP SEQUENCE OF 1-36 FROM N.A.  
 RX MEDLINE=85289338; PubMed=3161893;  
 RA Fisher R., Waller E.K., Grossi G., Thompson D., Tizard R.,  
 RA Schleuning W.-D.;  
 RT "Isolation and characterization of the human tissue-type plasminogen  
 activator structural gene including its 5' flanking region."; (1985).  
 RL J. Biol. Chem. 260:11223-11230 (1985).  
 RN [11]  
 RP SEQUENCE OF 31-562 FROM N.A.  
 RX MEDLINE=91291340; PubMed=1368681;  
 RA Itagaki Y., Yasuda H., Morinaga T., Mitsuoka S., Higashio K.;  
 RT "Purification and characterization of tissue plasminogen activator  
 secreted by human embryonic lung diploid fibroblasts, IMR-90 cells."; (1991).  
 RL Agric. Biol. Chem. 55:1225-1232 (1991).  
 RN [12]  
 RP SEQUENCE OF 36-562.  
 RC TISSUE=Melanoma;  
 RX MEDLINE=85000468; PubMed=6433976;  
 RA Pohl G., Kaelinstroem M., Bergsdorf N., Wallen P., Joernvall H.;  
 RT "Tissue plasminogen activator: peptide analyses confirm an indirectly  
 derived amino acid sequence, identify the active site serine residue,  
 establish glycosylation sites, and localize variant differences."; (1984).  
 RL Biochemistry 23:3701-3707 (1984).  
 RN [13]  
 RP SEQUENCE OF 33-52 AND 311-330.  
 RC TISSUE=Melanoma;  
 RX MEDLINE=83209620; PubMed=6682760;  
 RA Wallen P., Pohl G., Bergsdorf N., Raanby M., Ny T., Joernvall H.;  
 RT "Purification and characterization of a melanoma cell plasminogen  
 activator."; (1983).  
 RL Eur. J. Biochem. 132:681-686 (1983).  
 RN [14]  
 RP STRUCTURE OF CARBOHYDRATES.  
 RX MEDLINE=90092112; PubMed=2513186;  
 RA Pfeiffer G., Schmidt M., Strube K.-H., Geyer R.;  
 RT "Carbohydrate structure of recombinant human uterine tissue  
 plasminogen activator expressed in mouse epithelial cells."; (1989).  
 RL Eur. J. Biochem. 186:273-286 (1989).  
 RN [15]  
 RP CARBOHYDRATE-LINKAGE SITE THR-96.  
 RX MEDLINE=91159408; PubMed=1900431;  
 RA Harris R.J., Leonard C.K., Guzzetta A.W., Spellman M.W.;  
 RT "Tissue plasminogen activator has an O-linked fucose attached to  
 threonine-61 in the epidermal growth factor domain."; (1991).  
 RL Biochemistry 30:2311-2314 (1991).  
 RN [16]  
 RP DISULFIDE BONDS IN KRINGLE 2.  
 RX MEDLINE=91244765; PubMed=1645336;  
 RA Vlahos C.J., Wilhelm O.G., Hassell T., Jaskunas S.R., Bang N.U.;  
 RT "Disulfide pairing of the recombinant kringle-2 domain of tissue  
 plasminogen activator produced in *Escherichia coli*."; (1991).  
 RL J. Biol. Chem. 266:10070-10072 (1991).  
 RN [17]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF CATALYTIC DOMAIN.  
 RX MEDLINE=96200985; PubMed=8613982;  
 RA Lamba D., Bauer M., Huber R., Fischer S., Rudolph R., Kohnert U.,  
 RA Bode W.;  
 RT "The 2.3 Å crystal structure of the catalytic domain of recombinant  
 two-chain human tissue-type plasminogen activator."; (1996).  
 RL J. Mol. Biol. 258:117-135 (1996).  
 RN [18]  
 RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF CATALYTIC DOMAIN.  
 RX MEDLINE=97449126; PubMed=9305622;  
 RA Renatus M., Engh R.A., Stubbs M.T., Huber R., Fischer S., Kohnert U.,  
 RA Bode W.;  
 RT "Lysine 156 promotes the anomalous proenzyme activity of tPA: X-ray  
 crystal structure of single-chain human tPA."; (1997).



[1]  
 SEQUENCE FROM N.A.  
 TISSUE=Salivary gland;  
 MEDLINE=92039036; PubMed=1937019;  
 RA Kraetschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,  
 RA Alagon A., Donner P., Schleuning W.D.;  
 RT "The plasminogen activator family from the salivary gland of the  
 RT vampire bat *Desmodus rotundus*: cloning and expression.";  
 RL Gene 105:229-237(1991).  
 [2]  
 CHARACTERIZATION.  
 RP MEDLINE=93393059; PubMed=1309059;  
 RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,  
 RA Kraetschmar J., Haendler B., Langer G., Baldus B., Witt W.,  
 RA Donner P.;  
 RT "Plasminogen activators from the saliva of *Desmodus rotundus* (common  
 RT vampire bat): unique fibrin specificity.";  
 RL Ann. N.Y. Acad. Sci. 667:395-403(1992).  
 CC -!- FUNCTION: Probably essential to support the feeding habits of this  
 CC exclusively haematophagous animal. Probable potent thrombolytic  
 CC agent.  
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
 CC plasminogen to form plasmin.  
 CC -!- SUBUNIT: Monomer.  
 CC -!- SIMILARITY: Belongs to peptidase family S1.  
 CC -!- SIMILARITY: Contains 1 EGF-like domain.  
 CC -!- SIMILARITY: Contains 1 kringle domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M63989; AAA31594.1; -;  
 DR PIR; J50599; J50599  
 DR HSP; P98119; IAS1.  
 DR MEROPS; S01.239; -;  
 DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
 DR InterPro; IPR006209; EGF\_like.  
 DR InterPro; IPR006210; IEFG.  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR Pfam; PF00008; EGF; 1.  
 DR Pfam; PF00051; Kringle; 1.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PRO0722; CHYMOTRYPSIN.  
 DR PRINTS; PRO0018; KRINGLE.  
 DR ProDom; PD000395; Kringle; 1.  
 DR SMART; SM00181; EGF; 1.  
 DR SMART; SM00130; KR; 1.  
 DR SMART; SM00020; Tryp\_Spc; 1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS00026; EGF\_3; 1.  
 DR PROSITE; PS00021; KRINGLE\_1; 1.  
 DR PROSITE; PS00070; KRINGLE\_2; 1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
 KW Kringle; EGF-like domain; Signal; Multigene family.  
 FT SIGNAL 1 36 POTENTIAL.  
 FT CHAIN 37 431 SALIVARY PLASMINOGEN ACTIVATOR BETA.  
 FT DOMAIN 37 75 EGF-LIKE.  
 FT DOMAIN 82 163 KRINGLE.  
 FT DOMAIN 179 431 SERINE PROTEASE.  
 FT ACT\_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 275 275 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 382 382 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT DISULFID 41 52 BY SIMILARITY.  
 FT DISULFID 46 63 BY SIMILARITY.  
 FT DISULFID 65 74 BY SIMILARITY.  
 FT DISULFID 82 163 BY SIMILARITY.  
 FT DISULFID 103 145 BY SIMILARITY.  
 FT DISULFID 134 158 BY SIMILARITY.  
 FT DISULFID 168 299 BY SIMILARITY.  
 FT DISULFID 211 227 BY SIMILARITY.  
 FT DISULFID 219 288 BY SIMILARITY.  
 FT DISULFID 313 388 BY SIMILARITY.  
 FT DISULFID 345 361 BY SIMILARITY.  
 FT DISULFID 378 406 BY SIMILARITY.  
 FT CARBOHYD 139 139 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 431 AA; 48221 MW; 699B5675B162CBF CRC64;  
 Query Match 38.2%; Score 862.5; DB 1; Length 431;  
 Best Local Similarity 44.1%; Pred. No. 4.3e-62;  
 Matches 175; Conservative 59; Mismatches 146; Indels 17; Gaps 8;  
 QY 9 SNCDCLNGGTCVSNKYFSNIHCNCPKFGQGHCEIDKSKTCYEGNGHFYRGKASTDTMG 68  
 DB 42 SELRCFNGGTCWQAASFDF-VCOCPKGYTKQCEVDTHATCYKDGQGVYRGVSTSSG 100  
 QY 69 RPLCPWNSATVLQOYTHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLKPLVQECMV 128  
 DB 101 AQCINWNSNLTTRTYNGERSDAITLGLGNHNYCRNPDRNNSKPCYVIKASKFILEPCSV 160  
 QY 129 HDCADGKLKFCQG-QKTLRPFKIIGGEFTTIENQWPFAALYRRHRGS-VTVCGGSLI 186  
 DB 161 PVCS-----KATGLRKYPQLHSTGGFTDITSHFWQAALFAQNRSSGGERFLCGGILI 216  
 QY 187 SPCWISATHCFID-YPKKEDYIVLGRSLNSNTQGMKEVENLILHKDYASDTLAHH 245  
 DB 217 SSCWVITAAHCFQERYPPQHLLV-V-LGTYRVKPKESQTEVEKCIIEEFDDET--YN 273  
 QY 246 NDIALKIRSKGRCAQPSRTIOTICLPSMYNDPFGTSCITGKSKENSDYLYPEOLK 305  
 DB 274 NDIALQLKSGSPQCAQESDSVRAICLPEANLQLPDWTECELSGSGKSKSPFPYSEOLK 333  
 QY 306 MTWVKLIISHRECOQPHYGSEVTTKMLCAADPOWKT-----DSQCQSGGGLVCSLQQR 359  
 DB 334 EGHVRLYPSRSTSKFELFKNTVTNMLCAGDTSRSEIYFNVHDAQCQSGGGLVCMNDNH 393  
 QY 360 MTLTGIVSWGRGKALKDKPGVTVRYSHFLFWIRSHTK 396  
 DB 394 MTLGIISWGVGGEGKDPGVTVTKVTVLGVIRDNMR 430  
 RESULT 10  
 TPA\_HUMAN  
 ID TPA\_HUMAN STANDARD; PRT; 562 AA.  
 AC P00750; Q15103;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)  
 DE (t-PA) (t-plasminogen activator) (Alteplase) (Reteplase).  
 GN PLAT.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Melanoma;  
 RX MEDLINE=93115262; PubMed=6337343;  
 RA Pennica D., Holmes W.E., Kohr W.J., Harkins R.N., Vehar G.A.,  
 RA Ward C.A., Bennett W.F., Velverton E., Seeburg P.H., Heyneker H.L.,  
 RA Goeddel D.V., Collen D.;  
 RT "Cloning and expression of human tissue-type plasminogen activator  
 RT cDNA in *E. coli*.";  
 RL Nature 301:214-221(1983).



RL J. Biol. Chem. 264:17947-17952 (1989).  
 RP CHARACTERIZATION.  
 RX MEDLINE=9393059; PubMed=1309059;  
 RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,  
 RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,  
 RA Donner P.;  
 RT "Plasminogen activators from the saliva of *Desmodus rotundus* (common vampire bat): unique fibrin specificity.";  
 RL Ann. N.Y. Acad. Sci. 667:395-403 (1992).  
 CC -!- FUNCTION: Probably essential to support the feeding habits of this exclusively haematophagous animal. Probable potent thrombolytic agent.  
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.  
 CC -!- ENZYME REGULATION: Activity toward plasminogen is stimulated in the presence of fibrin I.  
 CC -!- SUBUNIT: Monomer.  
 CC -!- DOMAIN: The fibronectin type-I domain mediates binding to fibrin, and the kringle domain apparently mediates fibrin-induced stimulation of activity.  
 CC -!- SIMILARITY: Belongs to peptidase family S1.  
 CC -!- SIMILARITY: Contains 1 EGF-like domain.  
 CC -!- SIMILARITY: Contains 1 fibronectin type I domain.  
 CC -!- SIMILARITY: Contains 1 kringle domain.  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL; M63988; AAA31593.1; -;  
 DR EMBL; J05082; AAA31596.1; -;  
 DR PIR; A34369; A34369.  
 DR PIR; J05098; J05098.  
 DR HSP; P98119; 1A51.  
 DR MEROPS; S01.242; -;  
 DR InterPro; IPR009003; Cys Ser trypsin.  
 DR InterPro; IPR006209; EGF-like.  
 DR InterPro; IPR000093; Fibrinectn.  
 DR InterPro; IPR006210; IEGF.  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR001254; Peptidase S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR Pfam; PF00008; EGF; 1.  
 DR Pfam; PF00039; fn1; 1.  
 DR Pfam; PF00051; kringle; 1.  
 DR PRINTS; PF00089; trypsin; 1.  
 DR PRINTS; PF00722; CHYMOTRYPSIN.  
 DR PRINTS; PF00018; KRINGLE.  
 DR ProDom; PD000395; Kringle; 1.  
 DR SMART; SM00181; EGF; 1.  
 DR SMART; SM00058; FN1; 1.  
 DR SMART; SM00130; KR; 1.  
 DR SMART; SM00020; tryp\_Spc; 1.  
 DR PROSITE; PS00022; EGF 1; 1.  
 DR PROSITE; PS01186; EGF 2; 1.  
 DR PROSITE; PS00026; EGF 3; 1.  
 DR PROSITE; PS01253; FIBRONECTIN 1; 1.  
 DR PROSITE; PS00021; KRINGLE 1; 1.  
 DR PROSITE; PS00070; KRINGLE 2; 1.  
 DR PROSITE; PS02400; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 DR Plasminogen activation; Hydrolase; Serine protease; Glycoprotein; Kringle; EGF-like domain; Signal; Multigene family.  
 KW SIGNAL 1 36  
 FT CHAIN 37 477  
 FT DOMAIN 40 82  
 FT FIBRONECTIN TYPE-I.  
 FT EGF-LIKE.  
 FT

FT DOMAIN 128 209  
 FT ACT SITE 225 477  
 FT ACT SITE 272 272  
 FT ACT SITE 321 321  
 FT DISULFID 428 428  
 FT DISULFID 42 72  
 FT DISULFID 70 79  
 FT DISULFID 87 98  
 FT DISULFID 92 109  
 FT DISULFID 111 120  
 FT DISULFID 128 209  
 FT DISULFID 149 191  
 FT DISULFID 180 204  
 FT DISULFID 214 345  
 FT DISULFID 257 273  
 FT DISULFID 265 334  
 FT DISULFID 359 434  
 FT DISULFID 391 407  
 FT DISULFID 424 452  
 FT CARBOHYD 185 185  
 FT CARBOHYD 398 398  
 FT CONFLICT 403 403  
 FT CONFLICT 417 417  
 FT CONFLICT 435 435  
 FT CONFLICT 477 AA; 53719 MW; 17486555COB5077C CRC64;  
 SQ SEQUENCE 477 AA; 53719 MW; 17486555COB5077C CRC64;  
 Query Match 38.5%; Score 968.5; DB 1; Length 477;  
 Best Local Similarity 43.7%; Pred. No. 1.6e-62;  
 Matches 178; Conservative 60; Mismatches 148; Indels 21; Gaps  
 QY 3 ELHQVP-----SNCDLNGGTCVSNKYFSNIHWCNPKKFGGQHCIDKSTCYEGNGHY 58  
 DB 78 QCHTVFVKSCBELCFNGGTCQAAFSDF-VCCQPGYTGKQCEVDTHATCYKQGVY 136  
 QY 59 GKASTDTMGPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNPDNRRPWCYQVG 118  
 DB 137 RGTWSTSSGAQCINMNSNLTTRTYNGRSDAITLGLNHNHNYCRNPDNNSKWCYVKA 194  
 QY 119 LKPLVQECWHDGADGKLKFCQG-QKTLRPRFKLIIGSEFTIENQPFALYRHRGGS- 176  
 DB 197 SKFILEFCVPCVS-----KATCGLRKYKEPQLHSTGGTGLFDITSHMPQAAFAFQNRSSG 252  
 QY 177 VTYCGGSLISPCWISATHCFID-YPKGEDYIVYLGSRSLNSNTQGMKFVEVNLH 231  
 DB 253 ERFLCGGILISSCWLVTAACHFQERYPPQHLRVV-LGRTVYRKPGEQTFEVEKCI 311  
 QY 236 DYSADTLAHNDIALLLKIRSEKGCACQPSRTIQTICLPSMYNDPQFOTSCITFGKENS 296  
 DB 312 BFDDDT--YNNIDIALLOKSGSPQCSQESDVSRAICLPEANLQLPDWTECELSGYGKHS 365  
 QY 296 TDLYPEQLKMTVVVKLISHRECCQPHYGVSEVTTMLCAADPQWKT-----DSCQD 345  
 DB 370 SSPFYEQLEKGVRLVPSSRCTSKPLFNKTVTNMMLCAGDTRSGEYIPNVHDACQD 421  
 QY 350 GPLVCSLQGRMTLGIIVSGEGCALKDKPGVYTRVSHFLPKIRSHTK 396  
 DB 430 GPLVCMNDNMHTLLGIISWGVGCGEKDIPGVYTKVTNYLGNWIRDNM 476  
 RESULT 9  
 ID URTE DESRO STANDARD; PRT; 431 AA.  
 AC P98121;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DE 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Salivary plasminogen activator beta precursor (EC 3.4.21.68) (DSFA beta)  
 DE *Desmodus rotundus* (Vampire bat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;  
 OC Desmodontinae; *Desmodus*.  
 OX NCBI\_TaxID=9430;

28-FEB-2003 (Rel. 41, Last annotation update)  
 Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)  
 (U-plasminogen activator).  
 OC Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90110185; PubMed=2295632;  
 RA Leslie N.D., Kessler C.A., Bell S.M., Degen J.L.;  
 RT "The chicken urokinase-type plasminogen activator gene.";  
 RL J. Biol. Chem. 265:1339-1344(1990).  
 CC -|- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
 CC plasminogen to form plasmin.  
 CC -|- SIMILARITY: Belongs to peptidase family S1.  
 CC -|- SIMILARITY: Contains 1 EGF-like domain.  
 CC -|- SIMILARITY: Contains 1 kringle domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; J05187; AAA49131.1; -;  
 DR EMBL; J05188; AAA49130.1; -;  
 DR PIR; A35005; A35005.  
 DR HSP; P00763; IDPO.  
 DR MEROPS; S01.231; -;  
 DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
 DR InterPro; IPR006209; EGF-like.  
 DR InterPro; IPR006210; IEGF.  
 DR InterPro; IPR000001; Kringle.  
 DR InterPro; IPR008293; Pept\_S1A\_uPA.  
 DR InterPro; IPR001254; Peptidase\_S1.  
 DR InterPro; IPR001314; Peptidase\_S1A.  
 DR Pfam; PF00051; Kringle; 1.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PIRSF00144; Uro\_k\_plasm\_act; 1.  
 DR PRINTS; PRO0722; CHYMOTRYPSIN.  
 DR PRINTS; PRO0018; KRINGLE.  
 DR ProDom; PD000395; Kringle; 1.  
 DR SMART; SM00181; EGF; 1.  
 DR SMART; SM00130; KR; 1.  
 DR SMART; SM00020; Tryp\_Spc; 1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS00026; EGF\_3; 1.  
 DR PROSITE; PS00021; KRINGLE\_1; 1.  
 DR PROSITE; PS00070; KRINGLE\_2; 1.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
 KW Kringle; EGF-like domain; Signal; Zymogen.  
 FT SIGNAL 1 20 POTENTIAL.  
 FT CHAIN 21 434 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.  
 FT CHAIN 21 171 CHAIN A (BY SIMILARITY).  
 FT CHAIN 173 434 CHAIN B (BY SIMILARITY).  
 FT CHAIN 36 72 EGF-LIKE.  
 FT DOMAIN 79 158 KRINGLE.  
 FT DOMAIN 159 172 CONNECTING PEPTIDE.  
 FT DOMAIN 173 434 SERINE PROTEASE.  
 FT DISULFID 40 48 BY SIMILARITY.  
 FT DISULFID 42 60 BY SIMILARITY.  
 FT DISULFID 62 71 BY SIMILARITY.  
 FT DISULFID 162 296 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 202 218 BY SIMILARITY.  
 FT DISULFID 210 285 BY SIMILARITY.

FT DISULFID 310 379 BY SIMILARITY.  
 FT DISULFID 342 358 BY SIMILARITY.  
 FT DISULFID 369 397 BY SIMILARITY.  
 FT ACT\_SITE 217 217 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 272 272 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 373 373 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT CARBOHYD 228 228 N-LINKED (GLCNAC...). (POTENTIAL).  
 SQ SEQUENCE 434 AA; 49400 MW; BD881048DD666A55 CRC64;  
 Query Match 44.2%; Score 997; DB 1; Length 434;  
 Best Local Similarity 47.8%; Pred. No. 6.5e-73;  
 Matches 187; Conservative 64; Mismatches 122; Indels 18; Gaps 7;  
 QY 11 CDCINGGTCVSNKYPSNIHWKNCPKKGGQHCETDKSKTCYEGNGHYRGKASDITWGRP 70  
 DB 40 CQCLNGGTCITTFPSQIKRCLCEGYGLHCEIDTNSICISGNGEDYRGADDP---G 95  
 QY 71 CLPNSATVLO-QTYHAHRSDALQGLGKHNYCRPNRRPWCYVQVGLKPLVQE--CM 127  
 DB 96 CLYWDHPSVIRMGDYHADLKNALQGLGKHNYCRPNRGRPRWCYTK--RRYSIQETPCS 153  
 QY 128 VHDCAADGKLCQCGQKTLRPRFKIIGGFEFTIENQWFAIYRRHSGSVYVCGGSLIS 187  
 DB 154 TIE---KERTCGGRSFKYKIVGSGAEVETQPIAGIFQNM-GTQFLCGGSLID 208  
 QY 188 PCWISATHCFID----YPKKEDYIVYLGSRSLNSNTQGEKMFVENLILHKOYSADTLA 243  
 DB 209 PCWLTAACFCYNTKQPKNSVYKVFGLKSIILTNDEHEQVFMVDSIISHPDFTDTGG 268  
 QY 244 HENDIALLKIRSKGRCACQPSRTTQICLPSMYNDPOFGTSCETGFGKNSDLYLPEQ 303  
 DB 269 NNDIALIRIRITASQCAVESNYRTVCLPENLINDNTWCETIAGVGKQNSDIYYAOR 328  
 QY 304 LKMTVVYKLSHRECOQPHYGVSEVTVKMLCAADPQWKTDCQGSQGGPLVCSIGRMTLT 363  
 DB 329 LMSATVNLISQDDCNKYDSTRVTDNNMVCAGDELWETDACKGSDGSGPMVCEHNGRMTLY 388  
 QY 364 GIVSWGRCALKDKPGVYTRYVSHFLPWRSH 394  
 DB 389 GIVSWGDCGCKENKPGVYTRYVTRVTRVNLWDSN 419  
 RESULT 8  
 URT2 DESRO  
 ID URT2 DESRO STANDARD; PRT; 477 AA.  
 AC P15638;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Salivary plasminogen activator alpha 2 precursor (EC 3.4.21.68) (DSPA  
 DE alpha-2) (BAT-PA) (T-plasminogen activator).  
 OS Desmodus rotundus (Vampire bat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;  
 OC Desmodontinae; Desmodus.  
 OX NCBI\_TaxID=9430;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=Salivary gland;  
 RX MEDLINE=92039036; PubMed=1937019;  
 RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,  
 RA Alagon A., Donner P., Schleuning W.D.;  
 RT "The plasminogen activator family from the salivary gland of the  
 RT vampire bat Desmodus rotundus: cloning and expression.";  
 RL Gene 105:229-237(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX TISSUE=Salivary gland;  
 RX MEDLINE=90036867; PubMed=2509450;  
 RA Gardell S.J., Duong L.T., Diehl R.E., York J.D., Hare T.R.,  
 RA Register R.B., Jacobs J.W., Dixon R.A.F., Friedman P.A.;  
 RT "Isolation, characterization, and cDNA cloning of a vampire bat  
 RT salivary plasminogen activator.";

SMART; SMO0020; TYP SPC; 1.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
DR PROSITE; PS00026; EGF\_3; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS00070; KRINGLE\_2; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; FALSE\_NEG.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
KW Kringle; EGF-like domain; Zymogen; Signal.  
FT SIGNAL; 1 20 POTENTIAL.  
FT CHAIN; 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.  
FT CHAIN; 21 178 CHAIN A (BY SIMILARITY).  
FT CHAIN; 157 178 SHORT A CHAIN (A1).  
FT CHAIN; 180 433 CHAIN B (BY SIMILARITY).  
FT CHAIN; 28 64 EGF-LIKE.  
FT DOMAIN; 71 152 KRINGLE.  
FT DOMAIN; 153 179 CONNECTING PEPTIDE.  
FT DOMAIN; 180 433 SERINE PROTEASE.  
FT DISULFID; 32 40 BY SIMILARITY.  
FT DISULFID; 34 52 BY SIMILARITY.  
FT DISULFID; 54 63 BY SIMILARITY.  
FT DISULFID; 169 301 INTERCHAIN (BY SIMILARITY).  
FT DISULFID; 211 227 BY SIMILARITY.  
FT DISULFID; 219 290 BY SIMILARITY.  
FT DISULFID; 315 384 BY SIMILARITY.  
FT DISULFID; 347 363 BY SIMILARITY.  
FT DISULFID; 374 402 BY SIMILARITY.  
FT ACT\_SITE; 277 277 CHARGE RELAY SYSTEM.  
FT ACT\_SITE; 277 277 CHARGE RELAY SYSTEM.  
FT ACT\_SITE; 378 378 CHARGE RELAY SYSTEM.  
SQ SEQUENCE 433 AA; 48268 MW; A95C35F6250443P9 CRC64;  
Query Match 72.1%; Score 1626.5; DB 1; Length 433;  
Best Local Similarity 70.2%; Pred. No. 1.4e-123;  
Matches 283; Conservative 50; Mismatches 61; Indels 9; Gaps 2

QY 9 SNCDCLNGTGVSNKYFNSNIHWCNPKFGGHOCHIDSKTCYEGNGHFRGKASTDTMG 68  
Db 30 SNGCGQNGVCVSYKYFSRIKSCPRFQSEHCEIDASKTCYHNGSDYRGKANTDKG 89  
QY 69 RCLPNSATVLCQYTHAHRSDALQGLGKHNKCNPNRRPWCYVQGLKPLVQECMV 128  
Db 90 RCLAWNAVPLQKFNAPHRDAISLGLGKHNKCNPNDRPWCYVQGLRQFVQECMV 149  
QY 129 HDCADGKLGK-FCGOKTLRPRFKLIGGFTTIENQPFWFAALVRRHGGG-VTY 179  
Db 150 HDCSLSKPSSVDQGGQCGKALRPRFKVGGFTTEVENQPFWFAALVQKNGGSPSPF 209  
QY 180 VCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKPEVENILHKDYSA 239  
Db 210 KCGGSLISPCWVSAAHCFIQLPKKENVYVYLGSKSSYNPGENKFEVQILHEYYRE 269  
QY 240 DTLAHNDIALKIRSKGRCACQPSRTIOTICLPSWYNDPQGTSCETITGFKENSTDYL 299  
Db 270 DSLAYHNDIALKIRTSQCAQPSRSIQTICLPPRFTDAPGSDCEITGFKESDYL 329  
QY 300 YPEQLKNTVYVYVLIHRECOQPHYVGSSEVTTKMLCAADPQWKTSCQSGGGLVCSLQGR 359  
Db 330 YPKNLKMSVYVYVLIHRECOQPHYVGSSEVTTKMLCAADPQWKTSCQSGGGLVCSLQGR 389  
QY 360 MLTGVSVWGRGCAKDKKGVTVTRVSHFLPWIRSHRTKEENGLA 402  
Db 390 PTLGIVSVWGRGCAKDKKGVTVTRVSHFLPWIRSHRTKEENGLA 432

RESULT 7  
UROK\_CHICK  
ID UROK\_CHICK STANDARD; PRT; 434 AA.  
AC P15120;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)

QY 360 MLTGVSVWGRGCAKDKKGVTVTRVSHFLPWIRSHRTKEENGLA 402  
Db 389 PTLGIVSVWGRGCAKDKKGVTVTRVSHFLPWIRSHRTKEENGLA 431

RESULT 6  
UROK\_MOUSE  
ID UROK\_MOUSE STANDARD; PRT; 433 AA.  
AC P06859;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)  
GN PLAU.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85179474; PubMed=2985383;  
RA Belin D., Vassalli J.-D., Conbepine C., Godeau F., Nagamine Y.,  
RA Reich E., Kocher H.P., Duvoisin R.M.,  
RT "Cloning, nucleotide sequencing and expression of cDNAs encoding  
RT mouse urokinase-type plasminogen activator.";  
RL Eur. J. Biochem. 148:225-232(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98163499; PubMed=2831940;  
RA Degen S.J.F., Hecker J.L., Reich E., Degen J.L.;  
RA "The murine urokinase-type plasminogen activator gene.";  
RL Biochemistry 26:8270-8279(1987).  
CC -|- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
CC plasminogen to form plasmin.  
CC -|- SUBUNIT: Found in high and low molecular mass forms. Each consists  
CC of two chains, A and B. The high molecular mass form contains a  
CC long chain A. Cleavage occurs after residue 156 in the low  
CC molecular mass form to yield a short A1 chain (By similarity).  
CC -|- SIMILARITY: Belongs to peptidase family S1.  
CC -|- SIMILARITY: Contains 1 EGF-like domain.  
CC -|- SIMILARITY: Contains 1 kringle domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X02389; CAA26231.1; -;  
CC EMBL; M17922; AAA40539.1; -;  
CC PIR; A29420; UTKMS.  
CC HSSP; P00749; 1KDU.  
CC MEROPS; S01.231; -;  
CC MGD; MGI:97611; Plau.  
DR InterPro; IPR009003; Cys Ser trypsin.  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR006210; IEGF.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR008293; Pept S1a uPA.  
DR InterPro; IPR001254; Peptidase S1.  
DR InterPro; IPR001314; Peptidase\_S1a.  
DR Pfam; PF00051; kringle; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PIRSF; PIRSF001144; UroK\_PlasM\_act; 1.  
DR PRINTS; PRO0722; CHYMOTRYPSIN.  
DR PRINTS; PRO0018; KRINGLE.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00181; EGF; 1.  
DR SMART; SM00130; KR; 1.

DB 201 RHGGGSIYLCGGSLSPWVYSATCFIDHPKKNYIVYLQGRSLNSDTRGEMQFEVEK 260  
 QY 231 LILHKYSDATLAHNDIALIKIRSKGRCAPSRITQICLPSVMYNDPQFCTSCITGF 290  
 DB 261 LILHEDYSAESLAHNDIALIKIRSGQCAQSRSIQICLPPEHDAHSTREITGF 320  
 QY 291 GKENSTDYLYPEQLKMTVVKLISHRECQQPHYGVSEVTTKMLCAADPQWKTDSCQDSSG 350  
 DB 321 GKENPSDYRSLDKMTFVSLVSGHEVCQQPHYGAETDKMLCAADPQWKTDSCQDSSG 380  
 QY 351 PLVCSLQGRMTLGIYSWGGCALDKPGVYTVSHFLPWISHTKEENGLAL 403  
 DB 381 PLVCTIQGRULTLGIYSWGGRCANKYKPGVYTVSRKFLPWINTHTRGEINLVL 433

RESULT 5  
 UROK\_RAT STANDARD; PRT; 432 AA.  
 AC P29598;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)  
 DE (U-plasminogen activator).  
 GN PLAU.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Fischer 344;  
 RX MEDLINE=92233409; PubMed=1568219;  
 RA Henderson B.R.; Tansey W.P.; Phillips S.M.; Ramshaw I.A.,  
 RA Keiford R.F.;  
 RT "Transcriptional and posttranscriptional activation of urokinase  
 RT plasminogen activator gene expression in metastatic tumor cells";  
 RL Cancer Res. 52:2489-2496(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Rabbani S.A.;  
 RL Submitted (Apr-1992) to the EMBL/GenBank/DBJ databases.  
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg--Val bond in  
 CC plasminogen to form plasmin.  
 CC -!- SUBUNIT: Found in high and low molecular mass forms. Each consists  
 CC of two chains, A and B. The high molecular mass form contains a  
 CC long chain A. Cleavage occurs after residue 156 in the low  
 CC molecular mass form to yield a short A1 chain (By similarity).  
 CC -!- SIMILARITY: Belongs to peptidase family S1.  
 CC -!- SIMILARITY: Contains 1 EGF-like domain.  
 CC -!- SIMILARITY: Contains 1 kringle domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X63434; CAA45028.1; -;  
 CC EMBL; X65651; CAA46601.1; -;  
 CC PIR; S24604; S18932.  
 CC RSP; P00749; 1KDU.  
 CC MEROPS; S01.231; -;  
 CC InterPro; IPR009003; Cys Ser trypsin.  
 CC InterPro; IPR006209; EGF-like.  
 CC InterPro; IPR006210; IEGF.  
 CC InterPro; IPR000001; Kringle.  
 CC InterPro; IPR008293; Pept S1A uPA.  
 CC InterPro; IPR001254; Peptidase\_S1.

DR InterPro; IPR001314; Peptidase\_S1A.  
 DR Pfam; PF00051; Kringle; 1.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PIRSF; PIRSF001144; Uro plasmin act; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR PRINTS; PR00018; KRINGLE.  
 DR ProDom; PD000395; Kringle; 1.  
 DR SMART; SM00181; EGF; 1.  
 DR SMART; SM00130; KR; 1.  
 DR SMART; SM00020; Tryp\_Spc; 1.  
 DR PROSITE; PS00222; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
 DR PROSITE; PS00266; EGF\_3; 1.  
 DR PROSITE; PS00021; KRINGLE\_1; 1.  
 DR PROSITE; PS00070; KRINGLE\_2; 1.  
 DR PROSITE; PS02040; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; FALSE\_NEG.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
 KW Kringle; EGF-like domain; Zymogen; Signal.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT CHAIN 20 432 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.  
 FT CHAIN 20 177 CHAIN A (BY SIMILARITY).  
 FT CHAIN 156 177 SHORT A CHAIN (A1) (BY SIMILARITY).  
 FT CHAIN 179 432 CHAIN B (BY SIMILARITY).  
 FT DOMAIN 27 63 EGF-LIKE.  
 FT DOMAIN 70 151 KRINGLE.  
 FT DOMAIN 152 178 CONNECTING PEPTIDE.  
 FT DOMAIN 179 432 SERINE PROTEASE.  
 FT DISULFID 31 39 BY SIMILARITY.  
 FT DISULFID 33 51 BY SIMILARITY.  
 FT DISULFID 53 62 BY SIMILARITY.  
 FT DISULFID 168 300 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 210 226 BY SIMILARITY.  
 FT DISULFID 218 289 BY SIMILARITY.  
 FT DISULFID 314 383 BY SIMILARITY.  
 FT DISULFID 346 362 BY SIMILARITY.  
 FT DISULFID 373 401 BY SIMILARITY.  
 FT ACT\_SITE 225 276 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 276 276 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 377 377 CHARGE RELAY SYSTEM.  
 FT CONFLICT 16 16 N -> H (IN REF. 2).  
 FT CONFLICT 24 24 E -> G (IN REF. 2).  
 FT CONFLICT 32 32 D -> N (IN REF. 2).  
 FT SEQUENCE 432 AA; 47957 MW; 4EB1B96C716244C8 CRC64;  
 SQ

Query Match 73.4%; Score 1656.5; DB 1; Length 432;  
 Best Local Similarity 72.2%; Pred. No. 5.6e-126;  
 Matches 291; Conservative 45; Mismatches 58; Indels 9; Gaps 2;

QY 9 SNCDCLNGTCVSNKYEGNIHWCNCPKFGGQHCIDSKTCYEGNHGFRKASTDTWG 68  
 DB 29 SNGCGQGVGVSVYKFSIRRCSPKFKGHEIDTSKTCYHNGQSYRGKANTDKG 88  
 QY 69 RPLCPNSATVLQOTYHAHRSDALQLGLGKHNYCRPNRRPPWCYVQVGLKPLVQECMV 128  
 DB 89 RPLCAWNSPAVLQOTYNAHRSDALSLGLGKHNYCRPNDRPPWCYVQVGLKPLVQECMV 148  
 QY 129 HDCADGKLGK-----FCGGKTLRPPFKITGGFTTIENQWFAATVRRHRGGS-VTY 179  
 DB 149 QDCSLSKFSSSTVDQGGFCGQKALRPRFKIVGGFTTVVQVQVFAATVLRKNGSGPSF 208  
 QY 180 VCGGSLISPCWISATHCFIDYFKKEDYIVYLRSLNSNTGEMKFEVENILHKDYSA 239  
 DB 209 KCGSLISPCWASATHCFVQKKEEYVYVYLGQSKRNSYNGEMKFEVEQLIHEDFSD 268  
 QY 240 DTAHNDIALKIRSGRCAPSRITQICLPSMYNDPQFCTSCITGFGKENSTDYL 299  
 DB 269 ETIAFHNDIALKIRSGRCAPSRITQICLPPRFAPGSDCEITGFGQESATDYF 328  
 QY 300 YPEQLKMTVVKLISHRECQQPHYGVSEVTTKMLCAADPQWKTDSCQDSSGGLVCSLQGR 359  
 DB 329 YPKDLKSVVKLISHRECQQPHYGVSEVTKMLCAADPEWKTDSCQDSSGGLICNDGR 388

Db 21 SHELHOESGASNGCLNGKCVSYKFSNIQRCSPKPKFGHECHIDTSQTCEGNGHSY 80  
QY 59 RGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNKYNCRPNRRPWCYVQVG 118  
Db 81 RGNANTNTGRRPCLPWNSATVLLNTYHAHREDALQLGLGKHNKYNCRPNRRPWCYVQVG 140  
QY 119 LKPLVQECMVHDCADG-----KLKFCQCGOKTLPRPKIIGEBETIEN 161  
Db 141 LKOLVQECMVPCNGSGSHRPAVDGKNPSTPEKVEFCGQKALRPRFKIVGKSTIEN 200  
QY 162 QPFAAIYRRHRGGSVTVYCGGSLISPCWVWISATHCFIDYPPKEDYIVVLGRSLNSNTQ 221  
Db 201 QPFAAIYRRHRGGSVTVYCGGSLISPCWVWISATHCFINQQKEDYIVVLGRSLNSHT 260  
QY 222 GEMKFEVENILKHVSADTLAHHNDIALIKRSKRCQAQPSRTITICLSMNDPOF 281  
Db 261 GEMKFEVEKILHEDYSADSLAHHNDIALIKRTDQCAQPSRSQTICLFPVNGDAH 320  
QY 282 GTCSEITGFGKENSIDYLPQELKMTWKVLSHRECCQPHYGVSEVTTKMLCAADPQWKT 341  
Db 321 GASCEIVGFKNEPDIYLPQELKMTVKLVSHRECCQPHYGVSEVTTKMLCAADPQWKT 380  
QY 342 DSCQSGGGLVCSLQGMVTLTIVSNGRCALKKXGVVTVRVSHPLMIRSHTEENGL 401  
Db 381 DSCQSGGGLVCSLQGMVTLTIVSNGRCALKKXGVVTVRVSHPLMIRSHTEENGL 440  
QY 402 A 402  
Db 441 A 441

RESULT 4  
UROK\_BOVIN STANDARD; PRT; 433 AA.  
AC Q0558; Q2809;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Urokinase-type plasminogen activator precursor (BC 3.4.21.73) (uPA)  
DE (U-plasminogen activator).  
GN PLAÜ.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=aortic endothelium;  
RX MEDLINE=93216119; PubMed=8385052;  
RA Kraetzschmar J., Haendler B., Kojima S., Rifkin D.B.,  
RA Schleuning W.-D.;  
RT "Bovine urokinase-type plasminogen activator and its receptor:  
RT cloning and induction by retinoic acid.";  
RL Gene 125:177-183(1993).  
RN [2]  
RP SEQUENCE OF 12-433 FROM N.A.  
RC TISSUE=kidney;  
RT Ravn P., Berglund L., Petersen T.E.;  
RT "Cloning and characterization of the bovine plasminogen activators uPA  
RT and tPA.";  
RL Int. Dairy J. 5:605-617(1995).  
CC -|- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in  
CC plasminogen to form plasmin.  
CC -|- INDUCTION: By retinoic acid.  
CC -|- SIMILARITY: Belongs to peptidase family S1.  
CC -|- SIMILARITY: Contains 1 EGF-like domain.  
CC -|- SIMILARITY: Contains 1 kringle domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL; L03546; AAA51419.1; -;  
DR EMBL; X85801; CAA59796.1; -;  
DR PIR; JN0560; JN0560.  
DR HSSP; P00749; 1LMW.  
DR MEROPS; S01.231; -;  
DR InterPro; IPR009003; Cys Ser trypsin.  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR000001; Kringle.  
DR InterPro; IPR008293; Pept S1A uPA.  
DR InterPro; IPR001254; Peptidase S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00051; kringle; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR FIRSF; PIRSF001144; Utk plasm act; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR PRINTS; PR00018; KRINGLE.  
DR ProDom; PD000395; Kringle; 1.  
DR SMART; SM00130; KR; 1.  
DR SMART; SM00020; Tryp\_Spc; 1.  
DR PROSITE; PS00022; EGF\_1; 4.  
DR PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
DR PROSITE; PS00026; EGF\_3; 1.  
DR PROSITE; PS00021; KRINGLE\_1; 1.  
DR PROSITE; PS00070; KRINGLE\_2; 1.  
DR PROSITE; PS02040; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;  
KW Kringle; EGF-like domain; Signal; Zymogen.  
FT SIGNAL 1 20  
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.  
FT CHAIN 21 179 CHAIN A (BY SIMILARITY).  
FT CHAIN 181 433 CHAIN B (BY SIMILARITY).  
FT DOMAIN 29 65 EGF-LIKE.  
FT DOMAIN 72 153 KRINGLE.  
FT DOMAIN 154 180 CONNECTING PEPTIDE.  
FT DOMAIN 181 433 SERINE PROTEASE.  
FT DISULFID 33 41 BY SIMILARITY.  
FT DISULFID 35 53 BY SIMILARITY.  
FT DISULFID 55 64 BY SIMILARITY.  
FT DISULFID 170 201 INTERCHAIN (BY SIMILARITY).  
FT DISULFID 211 227 BY SIMILARITY.  
FT DISULFID 219 290 BY SIMILARITY.  
FT DISULFID 315 384 BY SIMILARITY.  
FT DISULFID 347 363 BY SIMILARITY.  
FT DISULFID 374 402 BY SIMILARITY.  
FT ACT\_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 277 277 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 378 378 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT CONFLICT 189 189 A -> T (IN REF. 2).  
SQ SEQUENCE 433 AA; 48730 MW; 4DE1B8D4DA47027A CRC64;

Query Match 76.6%; Score 1728; DB 1; Length 433;  
Best Local Similarity 73.8%; Pred. No. 9.8e-132;  
Matches 305; Conservative 44; Mismatches 54; Indels 10; Gaps 2  
QY 1 SHELHQV--PSNCDCLNGTCTVSNKYPSNIHWCNCPKFGGCHCIDSKTCYEGNGHY 58  
Db 21 SNEVHKESGESNCGCLNGKCVTYKFSNIQRCSPKPKFGHECHIDTSQTCEGNGHSY 80  
QY 59 RGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNKYNCRPNRRPWCYVQVG 118  
Db 81 RGNANRDLGRCLANDSPTVLKMYHAHRSALQLGLGKHNKYNCRPNRRPWCYVQVG 140  
QY 119 LKPLVQECMVHDCADGL-----KFCGQKTLRPRKIIGEBETIENQFWAAIYR 170  
Db 141 LKQFVQFCMVQDCSVGKSPSPREKEEFQCGQKALRPRFKIVGGQVTNAENQFWAAIYR 200  
QY 171 RHGGGSVTVYCGGSLISPCWVWISATHCFIDYPPKEDYIVVLGRSLNSNTQGMKFEVEN 230

```

FT DISULFID 52 61 BY SIMILARITY.
FT DISULFID 167 298 INTERCHAIN (BY SIMILARITY).
FT DISULFID 208 224 BY SIMILARITY.
FT DISULFID 216 287 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT ACT SITE 223 223 CHARGE RELAY SYSTEM.
FT ACT SITE 274 274 CHARGE RELAY SYSTEM.
FT ACT SITE 378 378 CHARGE RELAY SYSTEM.
FT CARBOHYD 324 324 N-LINKED (GLCNAC...) (BY SIMILARITY).
SQ SEQUENCE 433 AA; 48595 MW; 816D22DFEDC8792 CRC64;

Query Match 92.6%; Score 2090; DB 1; Length 433;
Best Local Similarity 90.6%; Pred. No. 7.3e-161;
Matches 375; Conservative 17; Mismatches 10; Indels 12; Gaps 3;

QY 1 SNELHVPSCDCLNGCTCVSNKYFNTHWNCNPKFGGHCETDKSKTCYEGNGHYRG 60
DB 21 SREL-QVPSDCGCLNGGTCMKNKYFSSLHWNCNPKFGGHCETDKSKTCYEGNGHYRG 79
QY 61 KASDTDMGRPLPNSATVLOQTYHAHRS DALQLGLGKHNYCRNPNRRRPWCYVQVGLK 120
DB 80 KASDTDMGRSCLAWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPNRRRPWCYVQVGLK 139
QY 121 PLVQECMWHDCADGK-----LKFQCGOKTAPRKLIGBEFTTIENQPWFAAIYRRH 172
DB 140 QRVQECMWHNCAADGKKSPPEELQFQCGQRTLRPRFKIVGGEFTTIENQPWFAAIYRRH 199
QY 173 RGSVTVYVCGSLSPCWVWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 232
DB 200 RGSVTVYVCGSLSPCWVWISATHCFINPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 259
QY 233 LHKYVSADTLAHNDIALKIRSKRGCAQPSRTIQTICLPSMYNDPQ---FGTSCEITG 289
DB 260 LHEDYSADTLAHNDIALKIRSKRGCAQPSRTIQTICLPSMYNDPNDPPFGTSCEITG 319
QY 290 FKENSTDYLYPEQLKMTVVKLVSHKQKQPHYGVSEVTTKMLCAADPWMTDSCQDGS 349
DB 320 FKENSTDYLYPEQLKMTVVKLVSHKQKQPHYGVSEVTTKMLCAADPWMTDSCQDGS 379
QY 350 GPLVCSLQGRMTLTGIVSGRCALKDKPGVTVRVSHFLPWIRSHTRKENGAL 403
DB 380 GPLVCSLQGHMTLTGIVSGRCALKDKPGVTVRVSHFLPWIRSHTRKENGAL 433

RESULT 3
UROK_PIG STANDARD; PRT; 442 AA.
AC P04185;
DT 20-MAR-1987 (Rel. 04, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)
DE (U-plasminogen activator).
GN PLAU.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=85087954; PubMed=6096832;
RA Nagamine Y., Pearson D., Altus M.S., Reich E.;
RT "cDNA and gene nucleotide sequence of porcine plasminogen activator.";
RL Nucleic Acids Res. 12:9525-9541(1984).
RN [2]
RP REVISION TO 241.
RA Nagamine Y.;
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.

```

```

CC -1- SIMILARITY: Belongs to peptidase family S1.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X01648; CAA25806.1; -.
CC EMBL: X02724; CAA26511.1; -.
CC PIR: A00932; UKPG.
CC HSP: P00749; LKDU.
CC MEROPS: S01.231; -.
CC InterPro: IPR009003; Cys Ser trypsin.
CC InterPro: IPR006209; EGF-like.
CC InterPro: IPR000001; Kringle.
CC InterPro: IPR008293; Pept_S1A_UPA.
CC InterPro: IPR001254; Peptidase_S1.
CC InterPro: IPR001314; Peptidase_S1A.
CC Pfam: PF00051; kringle; 1.
CC Pfam: PF00089; trypsin; 1.
CC PIRSF: PIRSF001144; Urk plasm act; 1.
CC PRINTS: PR00722; CHYNOTRYPsin.
CC PRINTS: PR00019; KRINGLE.
CC ProDom: PD000395; Kringle; 1.
CC SMART: SM00130; KR; 1.
CC SMART: SM00020; TRY_SPC; 1.
CC PROSITE: PS00022; EGF 1; 1.
CC PROSITE: PS01186; EGF 2; FALSE_NEG.
CC PROSITE: PS0026; EGF_3; 1.
CC PROSITE: PS00021; KRINGLE 1; 1.
CC PROSITE: PS00070; KRINGLE 2; 1.
CC PROSITE: PS00240; TRYPsin DOM; 1.
CC PROSITE: PS00134; TRYPsin HS; 1.
CC PROSITE: PS00135; TRYPsin SER; 1.
CC KX Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
CC Kringle; EGF-like domain; Zymogen; Signal.
CC SIGNAL 1 20 BY SIMILARITY.
FT CHAIN 21 442 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 188 CHAIN A (BY SIMILARITY).
FT CHAIN 190 442 CHAIN B (BY SIMILARITY).
FT DOMAIN 29 65 EGF-LIKE.
FT DOMAIN 72 153 KRINGLE.
FT DOMAIN 154 189 CONNECTING PEPTIDE.
FT DOMAIN 190 442 SERINE PROTEASE.
FT CARBOHYD 152 152 N-LINKED (GLCNAC... ).
FT DISULFID 33 41 BY SIMILARITY.
FT DISULFID 35 53 BY SIMILARITY.
FT DISULFID 55 64 BY SIMILARITY.
FT DISULFID 179 310 INTERCHAIN (BY SIMILARITY).
FT DISULFID 220 236 BY SIMILARITY.
FT DISULFID 228 299 BY SIMILARITY.
FT DISULFID 324 393 BY SIMILARITY.
FT DISULFID 356 372 BY SIMILARITY.
FT DISULFID 383 411 BY SIMILARITY.
FT ACT SITE 286 286 CHARGE RELAY SYSTEM.
FT ACT SITE 387 387 CHARGE RELAY SYSTEM.
FT ACT SITE 241 241 Q -> H (IN REF. 1; CAA25806).
FT CONFLICT 242 242 Q -> H (IN REF. 1; CAA26511).
FT CONFLICT 288 288 A -> GS (IN REF. 1; CAA25806).
SQ SEQUENCE 442 AA; 49116 MW; EE32FCBF501321EE CRC64;

Query Match 81.0%; Score 1827.5; DB 1; Length 442;
Best Local Similarity 78.1%; Pred. No. 9.9e-140;
Matches 329; Conservative 32; Mismatches 41; Indels 19; Gaps 2;

QY 1 SNEHGV-PSNCDCLNGCTCVSNKYFNTHWNCNPKFGGHCETDKSKTCYEGNGHY 58

```









GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 25, 2004, 14:43:40 ; Search time 15.5408 Seconds  
(without alignments)  
1350.274 Million cell updates/sec

Title: US-09-880-503-6  
Perfect score: 2257  
Sequence: 1 SNELHQVPSNCDCNGGTCV.....VSHFLPWIRSKTEBGLAL 403

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2240	99.2	431	1	UROK_HUMAN
2	2090	92.6	433	1	UROK_PAPY
3	1827.5	81.0	442	1	UROK_PIG
4	1728	76.6	433	1	UROK_BOVIN
5	1656.5	73.4	432	1	UROK_RAT
6	1626.5	72.1	433	1	UROK_MOUSE
7	997	44.2	434	1	UROK_CHICK
8	868.5	38.5	477	1	URT2_DESRO
9	862.5	38.2	431	1	URT2_DESRO
10	858.5	38.0	562	1	TPA_HUMAN
11	841.5	37.3	559	1	TPA_RAT
12	840.5	37.2	477	1	URT1_DESRO
13	824.5	36.5	559	1	TPA_MOUSE
14	814	36.1	566	1	TPA_BOVIN
15	756	33.5	394	1	URTG_DESRO
16	726.5	32.2	655	1	HGFA_HUMAN
17	716	31.7	653	1	HGFA_MOUSE
18	700.5	31.0	603	1	PA12_CAVPO
19	681	30.2	615	1	PA12_HUMAN
20	638	28.3	593	1	PA12_BOVIN
21	508.5	22.5	790	1	PLMN_PIG
22	505	22.4	812	1	PLMN_MOUSE
23	501	22.2	810	1	PLMN_MACMU
24	499	22.1	810	1	PLMN_HUMAN
25	498.5	22.1	333	1	PLMN_CANFA
26	493	21.8	343	1	PLMN_SHEEP
27	484.5	21.5	812	1	PLMN_BOVIN
28	478	21.2	4548	1	AFOA_HUMAN
29	477	21.1	1420	1	AFOA_MACMU
30	447.5	19.8	338	1	PLMN_HORSE
31	447	19.8	875	1	NETR_HUMAN
32	434	19.2	761	1	NETR_MOUSE
33	431	19.1	810	1	PLMN_ERIEU

34	425.5	18.9	418	1	HATT_HUMAN
35	423	18.7	811	1	TMS6_HUMAN
36	417	18.5	436	1	HEPS_MOUSE
37	415.5	18.4	811	1	TMS6_MOUSE
38	408.5	18.1	343	1	PSS8_HUMAN
39	407.5	18.1	638	1	KAL_MOUSE
40	406.5	18.0	455	1	TMS5_MOUSE
41	405	17.9	277	1	KLKD_HUMAN
42	404.5	17.9	417	1	HEPS_HUMAN
43	402	17.8	638	1	KAL_HUMAN
44	401.5	17.8	855	1	ST14_HUMAN
45	401.5	17.8	855	1	ST14_MOUSE
46	401	17.8	416	1	HEPS_RAT
47	395.5	17.5	248	1	TRY3_CHICK
48	394.5	17.5	243	1	TRY1_BOVIN
49	391	17.3	457	1	TMS5_HUMAN
50	390.5	17.3	247	1	TRY2_BOVIN
51	390	17.3	263	1	CTR2_CANFA
52	390	17.3	342	1	PSS8_RAT
53	389.5	17.3	244	1	KLK6_HUMAN
54	388.5	17.2	453	1	TMS3_MOUSE
55	387.5	17.2	407	1	PA7_BOVIN
56	386.5	17.1	638	1	KAL_RAT
57	384.5	17.0	311	1	TRYG_MOUSE
58	384	17.0	342	1	PSS8_MOUSE
59	383	17.0	461	1	PRTC_HUMAN
60	382	16.9	263	1	CTR8_RAT
61	382	16.9	728	1	HGF_HUMAN
62	380	16.8	269	1	EL2_PIG
63	379.5	16.8	304	1	TRY3_HUMAN
64	378.5	16.8	264	1	CTRL_HUMAN
65	377	16.7	263	1	CTR8_HUMAN
66	376.5	16.7	247	1	TRY1_HUMAN
67	376	16.7	321	1	TRYG_HUMAN
68	376	16.7	454	1	TMS3_HUMAN
69	376	16.7	456	1	PRTC_BOVIN
70	375	16.6	250	1	KLK6_HUMAN
71	375	16.6	422	1	DES1_HUMAN
72	374.5	16.6	261	1	KLK6_MOUSE
73	374.5	16.6	728	1	HGF_MOUSE
74	374.5	16.6	728	1	HGF_RAT
75	374	16.6	271	1	EL2_RAT

ALIGNMENTS

RESULT 1	UROK_HUMAN	STANDARD;	PRT;	431 AA.
ID	UROK_HUMAN	Q15844; Q16618; Q969W6;		
AC	P00749; Q15844; Q16618; Q969W6;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	20-MAR-1987 (Rel. 04, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)			
GN	PLAU.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RX	SEQUENCE FROM N.A.			
RP	MEDLINE=85215647; PubMed=2987867;			
RA	Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Blasi F.;			
RT	"The human urokinase-plasminogen activator gene and its promoter."			
RL	Nucleic Acids Res. 13:2759-2771(1985).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Holmes W.E., Pennica D., Blaber M., Rey M.W., Guenzler W.A.,			
RA	Steffens G.J., Heyneker H.L.;			
RT	"Cloning and expression of the gene for pro-urokinase in Escherichia coli."			

Search completed: May 25, 2004, 14:58:34  
Job time : 26.311 secs

A26823  
 pancreatic elastase II (EC 3.4.21.71) precursor - pig  
 N:Alternate names: pancratopeptidase E  
 C:Species: Sus scrofa domestica (domestic pig)  
 C>Date: 16-Aug-1988 #sequence\_revision 16-Aug-1988 #text\_change 22-Jun-1999  
 C:Accession: A26823  
 R:Kawashima, I.; Tani, T.; Shimoda, K.; Takiguchi, Y.  
 DNA 5, 163-172, 1987  
 A>Title: Characterization of pancreatic elastase II cDNAs: two elastase II mRNAs are expressed  
 A:Reference number: A90958; MUID:87217962; PMID:3646943  
 A:Accession: A26823  
 A:Molecule type: mRNA  
 A:Residues: 1-269 <KAW>  
 A:Cross-references: GB:M16441; PIDN:AAA31027.1; PID:g164442  
 A:Superfamily: trypsin; trypsin homology  
 C:Keywords: hydrolase; serine proteinase  
 F:1-16/Domain: signal sequence #status Predicted <SIG>  
 F:17-28/Domain: propeptide #status Predicted <PRO>  
 F:29-269/Product: elastase II #status Predicted <MAT>  
 F:29-262/Domain: trypsin homology <TRY>  
 F:73,121,216/Active site: His, Asp, Ser #status predicted

Query Match 16.8%; Score 380; DB 2; Length 269;  
 Best Local Similarity 34.1%; Pred. No. 3.6e-22;  
 Matches 88; Conservative 50; Mismatches 106; Indels 14; Gaps 8;

QY 140 CQKTLRPRF-KIIGFEFTT-ENQWFAAIYRRHGGSVTVVCGGSLISPCWVISATHCF 198  
 DB 17 CGLPANLQPLRVVGGEDARENSPQVSL-QYDSSGQWRHCTGTLVDQSWVLTAHCI 75  
 QY 199 IDPKKEDYIVLGRSLRNSNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRKEG 258  
 DB 76 ---SSRTRYRVVLGRSLRNSNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRKEG 130  
 QY 259 RCAQPSRTIQITCLPSMYNDPQFGTSCITGKGENSTDYLYPQLKMTVVKLSHRECO 318  
 DB 131 --VSLTDKIQGLPAAGTILPNVYCVVTGWR-LQNGASPILOQGLLVVDYATCS 187  
 QY 319 QHYHGVSEVTTKMLCAADPQWKTSCQDGGGLVLC-SLQGRMTLTGIVSWGR--GGALK 375  
 DB 188 KPGWGSTVKTNMI CAGG-DGISSCNGDSGGLPLCQAGQWQVHGIVSGSLGCGNY 246  
 QY 376 DKPGVYTRVSHFLPIRS 393  
 DB 247 HKPSVTRVSNVIDWINS 264

RESULT 49  
 S33496  
 trypsin (EC 3.4.21.4) IV form a - human  
 C:Species: Homo sapiens (man)  
 C>Date: 03-Mar-1994 #sequence\_revision 03-Aug-1995 #text\_change 15-Aug-1997  
 C:Accession: S33496  
 R:Wiegand, U.; Corbach, S.; Minn, A.; Kang, J.; Mueller-Hill, B.  
 submitted to the EMBL Data Library, March 1993  
 A:Description: Identification, cloning and characterization of a cDNA encoding a human b  
 A:Reference number: S33496  
 A:Accession: S33496  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-304 <WIE>  
 A:Cross-references: EMBL:X72781  
 C:Genetics:  
 A:Gene: GDB:PRSS4; TRY4  
 A:Cross-references: GDB:335300  
 A:Map position: 7q35-7q35  
 C:Superfamily: trypsin; trypsin homology  
 C:Keywords: hydrolase; serine proteinase  
 F:81-296/Domain: trypsin homology <TRY>  
 F:120,164,257/Active site: His, Asp, Ser #status predicted

Query Match 16.8%; Score 379.5; DB 2; Length 304;  
 Best Local Similarity 37.7%; Pred. No. 4.5e-22;

Matches 92; Conservative 38; Mismatches 87; Indels 27; Gaps 9;

QY 150 KIIGFEFTTINQ-PWFAAIYRRHGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDIY 208  
 DB 80 KIVGG-YTCEENSLEFYQVSL-----NSGSHFCGGSLSISEQWVWSAAHCY-----KTRIQ 127  
 QY 209 VYLGRSLRNSNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRKEGCAQPSRTIQ 268  
 DB 128 VVLGHNITKVLGNEQFQINAAKIRPKYKRDTL--DNDIMLKLSSP----AVINARVS 181  
 QY 269 TITCLPSMYNDPQFGTSCITGKGENSTDYLYPQLKMTVVKLSHRECOQHYHGVSEVT 328  
 DB 182 TITSLPTA--PPAAGTECLISGNGNTLSPGADYDDELKCLDAFVLTQAECKAS--YPPGKIT 237  
 QY 329 TKMLCAADPQWKTSCQDGGGLVLC-SLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFL 388  
 DB 238 NSMFCVGFLEGGKDCQDGGFVVCNQ-----LQGVSWHGCAWKRPQGYTKVYNYV 293  
 QY 389 PMIR 392  
 DB 294 DMIK 297

RESULT 50  
 I38136  
 chymotrypsin-like proteinase (EC 3.4.21.-) CTRL-1 - human  
 C:Species: Homo sapiens (man)  
 C>Date: 06-Sep-1996 #sequence\_revision 06-Sep-1996 #text\_change 22-Jun-1999  
 C:Accession: I38136  
 R:Larsen, F.; Solheim, J.; Kristensen, T.; Kolsto, A.B.; Prydz, H.  
 Hum. Mol. Genet. 2, 1589-1595, 1993  
 A:Title: A tight cluster of five unrelated human genes on chromosome 16q22.1.  
 A:Reference number: I38135; MUID:94093544; PMID:8268911  
 A:Accession: I38136  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-264 <RES>  
 A:Cross-references: EMBL:X71874; NID:g406226; PIDN:CAA50710.1; PID:g406228  
 C:Genetics:  
 A:Gene: GDB:CTRL  
 A:Cross-references: GDB:204061  
 A:Map position: 16q22.1-16q22.1  
 A:Introns: 18/1; 52/3; 79/2; 106/3; 167/1; 211/3  
 C:Superfamily: trypsin; trypsin homology  
 C:Keywords: hydrolase; serine proteinase  
 F:34-257/Domain: trypsin homology <TRY>  
 F:75,121,214/Active site: His, Asp, Ser #status predicted

Query Match 16.8%; Score 378.5; DB 2; Length 264;  
 Best Local Similarity 35.0%; Pred. No. 4.6e-22;  
 Matches 90; Conservative 43; Mismatches 101; Indels 23; Gaps 10;

QY 140 CQKTLRPRF-----KIIGFEFTTINQWFAAIYRRHGGSVTVVCGGSLISPCWVISAT 195  
 DB 19 CGIPALKPALSPQRIVNGENAVLGSWPQVSL---QDSSGFHFCGGSLSISQSWVTTAA 74  
 QY 196 HCFIDYPKKEDYIVLGRSLRNSNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRS 255  
 DB 75 HCNVS-PGR--HPVVLGEYDRSSNAEPLQVLSVSRATHPSWNSITM--NNDVTLIKLAS 129  
 QY 256 KEGRCAQPSRTIQITCLPSMYNDPQFGTSCITGKGENSTDYLYPQLKMTVVKLSHR 315  
 DB 130 P-----AQYTRISPVCLASNEALTEGLTCVVTGRLSGVGNVTPAHLQQVALPLVTYN 185  
 QY 316 ECOQPHYGVSEVTTKMLCAADPQWKTSCQDGGGLVLC-SLQGRMTLTGIVSWG-RGCAL 374  
 DB 186 QCRQ--YWGSSITDSMICAGGA--GASSCGGSGGLVLCQKGNVTWVLIGVSGTNCNV 241  
 QY 375 KDKPGVYTRVSHFLPMI 391  
 DB 242 R-APAVYTRVSKPFTWI 257

R; Seki, T.; Ihara, I.; Sugimura, A.; Shimonishi, M.; Nishizawa, T.; Asami, O.; Hagiya, M. Biochem. Biophys. Res. Commun. 172, 321-327, 1990  
 A:Title: Isolation and expression of cDNA for different forms of hepatocyte growth factor  
 A:Reference number: A36677; MUID:91025062; PMID:2145836  
 A:Accession: B36677  
 A:Molecule type: mRNA  
 A:Residues: 1-728 <SE3>  
 A:Cross-references: GB:M60718; NID:g184031; PIDN:AAA52648.1; PID:g184032  
 A:Accession: A36677  
 A:Molecule type: mRNA  
 A:Residues: 1-161,167-728 <SE4>  
 A:Cross-references: EMBL:X16323  
 A:Experimental source: Leukocyte  
 R; Miyazawa, K.; Teubouchi, H.; Naka, D.; Takahashi, K.; Okigaki, M.; Arakaki, N.; Nakayama, Y. Biochem. Biophys. Res. Commun. 163, 967-973, 1999  
 A:Title: Molecular cloning and sequence analysis of cDNA for human hepatocyte growth factor  
 A:Reference number: A33512; MUID:89392017; PMID:2528952  
 A:Accession: A33512  
 A>Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-728 <MIY>  
 A:Cross-references: GB:M29145; NID:g184041; PIDN:AAA52650.1; PID:g306846  
 R; Rubin, J.S.; Chan, A.M.L.; Bottaro, D.P.; Burgess, W.H.; Taylor, W.G.; Cech, A.C.; Hlin, Proc. Natl. Acad. Sci. U.S.A. 88, 415-419, 1991  
 A:Title: A broad-spectrum human lung fibroblast-derived mitogen is a variant of hepatocyte growth factor  
 A:Reference number: A39006; MUID:9110540; PMID:1824873  
 A:Accession: A39006  
 A:Molecule type: mRNA  
 A:Residues: 1-161,167-728 <RUB>  
 A:Cross-references: GB:M55379  
 A:Experimental source: embryonic lung  
 R; Yoshiyama, Y.; Arakaki, N.; Naka, D.; Takahashi, K.; Hirono, S.; Kondo, J.; Nakayama, Y. Biochem. Biophys. Res. Commun. 175, 660-667, 1991  
 A:Title: Identification of the N-terminal residue of the heavy chain of both native and recombinant human hepatocyte growth factor  
 A:Reference number: PH0114; MUID:91207365; PMID:1826837  
 A:Accession: PH0114  
 A:Molecule type: Protein  
 A:Residues: 32-43,53-58 <YOS>  
 A:Experimental source: Plasma  
 R; Weidner, K.M.; Behrens, J.; Vandekekerckhove, J.; Birchmeier, W. J. Cell Biol. 111, 2097-2108, 1990  
 A:Title: Scatter factor: molecular characteristics and effect on the invasiveness of epithelial cells  
 A:Reference number: A37796; MUID:91035621; PMID:2146276  
 A:Accession: A37796  
 A:Molecule type: Protein  
 A:Residues: 86-91,329-344;356-363,'XX',366-370;425-434,442-447,'X',449-450;543-546,'X',547-548  
 R; Nakamura, T.; Nishizawa, T.; Hagiya, M.; Seki, T.; Shimonishi, M.; Sugimura, A.; Tashir, Nature 342, 440-443, 1989  
 A:Title: Molecular cloning and expression of human hepatocyte growth factor.  
 A:Reference number: S06794; MUID:90066676; PMID:2531289  
 A:Accession: S06794  
 A:Molecule type: mRNA  
 A:Residues: 1-31,'HK',34-77,'N',79-292,'V',294-299,'M',301-316,'A',318-335,'K',337-386,'X',387-388  
 A:Cross-references: EMBL:X16323; NID:g32081; PIDN:CAA4387.1; PID:g32082  
 A:Experimental source: liver  
 A:Note: The authors translated the codon CAG for residue 727 as Glu  
 R; Hartmann, G.; Naldini, L.; Weidner, K.M.; Sachs, M.; Vigna, E.; Comoglio, P.M.; Birchmeier, Proc. Natl. Acad. Sci. U.S.A. 89, 11574-11578, 1992  
 A:Title: A functional domain in the heavy chain of scatter factor/hepatocyte growth factor  
 A:Reference number: 159214; MUID:93087571; PMID:1260830  
 A:Accession: 159214  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-288,'ET',<HAR>  
 A:Cross-references: GB:L02931; NID:g184033; PIDN:AAA52649.1; PID:g184034  
 R; Miyazawa, K.; Kitamura, A.; Naka, D.; Kitamura, N. Eur. J. Biochem. 197, 15-22, 1991  
 A:Title: An alternatively processed mRNA generated from human hepatocyte growth factor gene  
 A:Reference number: S15443; MUID:91200041; PMID:1826653  
 A:Accession: S15443  
 A>Status: preliminary  
 A:Molecule type: mRNA

A:Residues: 1-288,'ET',<MIY>  
 A:Cross-references: EMBL:X57574; NID:g32083; PIDN:CAA40802.1; PID:g32084  
 R; Shima, N.; Nagao, M.; Ogaki, F.; Tsuda, E.; Murakami, A.; Higashio, K. Biochem. Biophys. Res. Commun. 180, 1151-1158, 1991  
 A:Title: Tumor cytotoxic factor/hepatocyte growth factor from human fibroblast  
 A:Reference number: 152253; MUID:92062058; PMID:1835383  
 A:Accession: 152253  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 161-166 <SHI>  
 A:Cross-references: GB:S62561; NID:g237996; PIDN:AAB20169.1; PID:g237997  
 C:Genetics:  
 A:Gene: GDB:HGF  
 A:Cross-references: GDB:127524; OMIM:142409  
 A:Map position: 7q21.1-7q21.1  
 A:Exons: 30/1; 85/2; 123/1; 161/2; 209/1; 249/2; 289/1; 347/2; 390/1; 424/2; 569/1;  
 C:Complex: disulfide-bonded heterodimer of chains derived from the same precursor  
 C:Function:  
 A:Description: stimulates mitosis of hepatocytes and other cells  
 A:Note: does not have proteinase activity  
 C:Superfamily: hepatocyte growth factor; kringle homology; trypsin homology  
 C:Keywords: alternative splicing; glycoprotein; growth factor; heterodimer; kringle  
 F:1-31/Domain: signal sequence #status predicted <SIG>  
 F:32-494/Domain: alpha chain #status experimental <MAT>  
 F:32-494/Domain: alpha chain #status experimental <ACH>  
 F:128-206/Domain: kringle homology <KR1>  
 F:211-288/Domain: kringle homology <KR2>  
 F:305-383/Domain: kringle homology <KR3>  
 F:391-469/Domain: kringle homology <KR4>  
 F:495-728/Domain: beta chain #status experimental <BCH>  
 F:495-728/Domain: trypsin homology <TRY>  
 F:32/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimental  
 F:294,402,566,633/Binding site: carboxylate (Asn) (covalent) #status predicted  
 F:487-604/Disulfide bonds: #status predicted  
 Query Match 16.9%; Score 382; DB 1; Length 728;  
 Best Local Similarity 26.9%; Pred. No. 7,4e-22;  
 Matches 114; Conservative 62; Mismatches 172; Indels 76; Gaps 19  
 QY 2 NELHQVPSNCDC--LNGGTCVSNKYFSNIHWC-----NCPKFGGQ--HCEIDSKTCYE 52  
 DB 335 HEHMDTPFNFKCKDLRENYC-RNPDGSESPWCFTDPNIRVGYCSQIFNCDSMGQDCYR 392  
 QY 53 GNGHFVRGKASTDTMGRPCLPWNSATVLQQTVAH---RSDALQLGLGKHNYCRNP-DNR 108  
 DB 394 GNGKYNMGNLSQTRSLGTCSMWDKN---MEDLHRHIFWEPDASKL---NENYCRNPDGDA 447  
 QY 109 RRPWCYVQVGLKPL----VQEC-----WVHDCADGKLFQCGQKTLPRFKI 151  
 DB 448 HGPWCYTNPLIPWDYCPISRCGDTTPTIVNLDPVISCATK-----QLRV 495  
 QY 152 IGGEFTTIENQWFAAIVRRHRGGSVTVVCGSLISPCWVISATHCFIDYPKKE--DYIV 205  
 DB 496 VNG-IPTRNIGWVSLRYRANK-----HICGSLIKESVLTARQCF---PSRLDKDYEA 546  
 QY 210 YLGRSLRNSNTQGMK--FEVENLILHKDYSADTLAHNDIALKIRKSGECAQPSRTI 267  
 DB 547 WLGIHDVHGRGDEKCKQVLNVSQLVYGEPS-----DLVLMKL-----ARPAVLDDFV 594  
 QY 268 QIICLPSVYNDQFQFTSCIEITGFGKENSVDYLPQLKMTVVKLISHRECQPHYVGSV 327  
 DB 595 STIDLPNYGCTTPEKTSQSVYNGWTGLIN--YDGLLRVAHYITMGNEKCSQHKGKVTL 652  
 QY 328 TTKMLCAADPQWKTDSQQDSGLVCSLQGRMTITGIVSGRGKALCKPKGVYTRVSHF 387  
 DB 653 NESEICAGAEKISGSGCEGDYGLPVCQHKQWMLGVIVPGRGCAIPNRPGIFVRVAYY 712  
 QY 388 LPWI 391  
 DB 713 AKWI 716  
 RESULT 48

A:Reference number: A44606; MUID:92184750; PMID:1544894  
A:Contents: annotation; beta-hydroxyaspartic acid  
C:Comment: Protein C is the zymogen of the vitamin K-dependent serine proteinase that in  
activation of factor Va is strongly enhanced by complexing with protein S. Protein C also  
bin, which cleaves a dodecapeptide from the amino end of the heavy chain; this reaction,  
C:Genetics:  
A:Gene: GDB:PLOC  
A:Cross-references: GDB:120317; OMIM:176860  
A:Map position: 2q13-2q21  
A:Introns: 24/1; 79/3; 88/1; 134/1; 179/1; 226/3; 266/1  
C:Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homology  
C:Keywords: anticoagulant; beta-hydroxyaspartic acid; blood coagulation; calcium binding  
F:1-32/Domain: signal sequence #status predicted <SIG>  
F:27-86/Domain: Gla domain homology <GUA>  
F:33-42/Domain: propeptide #status predicted <PRO>  
F:43-197/Product: protein C light chain #status predicted <LCH>  
F:92-131/Domain: EGF homology <EG1>  
F:140-175/Domain: EGF homology <EG2>  
F:200-461/Product: protein C heavy chain #status predicted <HCH>  
F:200-211/Domain: activation peptide #status experimental <APT>  
F:212-445/Domain: trypsin homology <TRY>  
F:48-49,56,58,61,62,67,68,71/Modified site: gamma-carboxyglutamic acid (Glu) #status exp  
F:59-64,92-105,101-120,122-131,140-151,147-160,162-175,183-319,238-254,373-387,398-426/H  
F:106-111/Disulfide bonds: #status predicted  
F:110/Binding site: carboxylate (Thr) (covalent) #status absent  
F:113/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status experimental  
F:139,290,355/Binding site: carboxylate (Asn) (covalent) #status experimental  
F:211-212/Cleavage site: Arg-Leu (thrombin) #status experimental  
F:253,299,402/Active site: His, Asp, Ser #status predicted  
F:371/Binding site: carboxylate (Asn) (covalent) (partial) #status atypical

Query Match 17.0%; Score 383; DB 1; Length 461;  
Best Local Similarity 28.2%; Pred. No. 3.8e-22;  
Matches 118; Conservative 50; Mismatches 147; Indels 104; Gaps 17;

QY 13 CLNGGTGCVSN-KYFNIHWNCNPKFGGQHCIEDKS-KTCYEGNGHFRGRKASTDTWGRP 70  
DB 105 CGGHTCNDIGSFS-----CDCRSGWGRFCQREVSFLNCLDNG----- 145  
QY 71 CLPWSATVLQTYHAHRSALQLGLGKHNYCRNPDNRFPWCYVQGLKPLVEQCMVHD 130  
DB 146 -----GCTHYCLEEVGWRGSCAPGYKLGDDLLQC--HP 177  
QY 131 CADGKLKFCQGGKTLRPRF-----IIGGEFTTIENQWFAAIVRRH 172  
DB 178 A-----VKFPCGRFWRMEKRSHLKRDTEDQDQVDPRLIDGKMTRRGDSFWQVYLLDSK 233  
QY 173 RGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYVYLGSRSLNSNTQGMKFEVENLI 232  
DB 234 K-----KLACGAVLIHPSWVLTAAHC-MDESKK-LLVRLGEYDLRWEKWLDDIDKEVP 286  
QY 233 LHKDVSADTLAHNDIALKIRSKGRCAQP---SRTTQICLPSM-----YNDPQGT 283  
DB 287 VHPNYSKSTT--DNDIALHL-----AQPATLSQTTVPICLPDSGLAERELN--QAQG 335  
QY 284 SCEITGFGKENSVDYLPQ-----LKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQW 339  
DB 336 ETLVTGWGHSREKAKNRFTVLFKIPVPHNECE--VMGNWVSENMLCAGILGD 393  
QY 340 KTDSCGDSGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVVYTRYSHFLPWRSHKTEE 398  
DB 394 RQDACGDSGGPVAASFHTWFLVGLVSWGEGGLLHNYGVYTKVSRVLDLTHGIRDK 452

RESULT 46  
KXRTB  
chymotrypsin (EC 3.4.21.1) B precursor - rat  
N:Alternate names: chymotrypsinogen B  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 18-Jun-1999  
C:Accession: A22658  
R:Bell, G.I.; Quinto, C.; Quiroga, M.; Valenzuela, P.; Craik, C.S.; Rutter, W.J.

J. Biol. Chem. 259, 14265-14270, 1984  
A:Title: Isolation and sequence of a rat chymotrypsin B gene.  
A:Reference number: A22658; MUID:85054881; PMID:6209274  
A:Accession: A22658  
A:Molecule type: DNA  
A:Residues: 1-263 <BE>  
A:Cross-references: GB:K02298; NID:g203653; PIDN:AAA98732.1; PID:g203654  
C:Genetics:  
A:Introns: 18/1; 52/3; 79/2; 105/3; 166/1; 210/3  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-33/Domain: propeptide #status predicted <PRO>  
F:34-263/Product: chymotrypsin B #status predicted <MAT>  
F:34-256/Domain: trypsin homology <TRY>  
F:35,120,213/Active site: His, Asp, Ser #status predicted

Query Match 16.9%; Score 382; DB 1; Length 263;  
Best Local Similarity 33.6%; Pred. No. 2.4e-22;  
Matches 87; Conservative 47; Mismatches 103; Indels 22; Gaps 7;

QY 138 FCGGKTLRPRF-----KIIGGEFTTIENQWFAAIVRRHSGSVTVYCGSLISPCWVIS 193  
DB 17 FCGGVPTIQVLTGLSRVINGEDALPGSWPQVSLQDK-----TGFHFCGSLISDWNVT 72  
QY 194 ATFCFIDYPKEDYIVYLGSRSLNSNTQGMKFEVENILHKDYSADTLAHNDIALKX 253  
DB 73 AAHGGV---KTSDDVVAAGEFPGGSDDEENIQVLKIAQVFNKPKNNFTV--RNDITLLKL 126  
QY 254 RSKEGRCAQPSITQICLPSMYNDPQGTSCIEITGFGKENSVDLYPEQLKMTVVKLIS 313  
DB 127 ATP-----AQSFTVSACVLPNVDDPPPGTVCATTGWTGKTKYNAUKTEKLAQALPIVS 182  
QY 314 HRBCQOPHYGVSEVTTKMLCAADPQWKTDCSGDSGGPLVCSLQGRMTLTGIVSWGRC 373  
DB 183 EADCKES--WGSKITDVTMCAGAS--GVSSCGDSGGPLVCQKQGVWTLAGIVSWGSGVC 238  
QY 374 LKDKPQVYTRYSHFLPWIR 392  
DB 239 STSTPAVYSRVTALMPWVQ 257

RESULT 47  
JH0579  
hepatocyte growth factor precursor [validated] - human  
N:Alternate names: hepatopoietin A; scatter factor  
C:Species: Homo sapiens (man)  
C:Date: 17-Aug-1992 #sequence\_revision 17-Aug-1992 #text\_change 08-Dec-2000  
C:Accession: JH0579; JH0333; A41140; B36677; A36677; A33512; A39006; PH0114; A3  
R:Seki, T.; Hagiya, M.; Shimonishi, M.; Nakamura, T.; Shimizu, S.  
Gene 102, 213-219, 1991  
A:Title: Organization of the human hepatocyte growth factor-encoding gene.  
A:Reference number: JH0579; MUID:91340155; PMID:1831432  
A:Accession: JH0579  
A:Molecule type: DNA  
A:Residues: 1-728 <SEK>  
A:Cross-references: DDBJ:D90318  
A:Note: the authors translated the codon GAA for residue 662 as Gly  
R:Seki, T.; Hagiya, M.; Shimonishi, M.; Nakamura, T.; Shimizu, S.  
submitted to JIPID, March 1991  
A:Description: Organization of the human hepatocyte growth factor-encoding gene  
A:Reference number: JH0333  
A:Accession: JH0333  
A:Molecule type: DNA  
A:Residues: 1-481, 'RT', 484-728 <SE2>  
R:Weidner, K.M.; Arakaki, N.; Hartmann, G.; Vandekerckhove, J.; Weingart, S.; E  
Proc. Natl. Acad. Sci. U.S.A. 88, 7001-7005, 1991  
A:Title: Evidence for the identity of human scatter factor and human hepatocyte  
A:Reference number: A41140; MUID:91334393; PMID:1831266  
A:Accession: A41140  
A:Molecule type: mRNA  
A:Residues: 1-728 <WFI>  
A:Cross-references: GB:M73239; NID:g337935; PIDN:AAA64239.1; PID:g337936

Db 282 SGWQLLBERG---VTARKLVVLPRLLTQDCLQQSQRFEG--GPVVTDNMFAGSYDGS 336

Qy 341 TDSQCGSGPLVCSLQGRWTLTGIVSWGCGALKDQPGVYTVSHPLPWIR---SHTKE 397

Db 337 KDACKGSGGPHAFRFGTFLTGWVSWGCGAAGHFGIYTVSRYTAWLRQLMGHPPS 396

Qy 398 ENG 400

Db 397 RQG 399

RESULT 44

KQRTPL

plasma kallikrein (EC 3.4.21.34) precursor - rat

N;Alternate names: Fletcher factor; kininogenin; serum kallikrein

C;Species: Rattus norvegicus (Norway rat)

C;Date: 30-Sep-1992 #sequence revision 30-Sep-1992 #text\_change 18-Jun-1999

A;Accession: A39180; A33320; S66851; I53041; S66852

R;Baubien, G.; Rosinski-Chupin, I.; Mattei, M.G.; Mbikay, M.; Chretien, M.; Seidah, N.G.

Biochemistry 30, 1628-1635, 1991

A;Title: Gene structure and chromosomal localization of plasma kallikrein.

A;Reference number: A39180; MUID:91129236; PMID:1993180

A;Accession: A39180

A;Molecule type: DNA

A;Residues: 1-638 <BEA>

A;Cross-references: GB:J05315

A;Note: the authors translated the codon GAG for residue 81 as Glu

R;Seidah, N.G.; Ladenheim, R.; Mbikay, M.; Hamelin, J.; Lutfalla, G.; Rougeon, F.; Lazur

DNA 8, 563-574, 1989

A;Title: The cDNA structure of rat plasma kallikrein.

A;Reference number: A33320; MUID:90091743; PMID:2598771

A;Accession: A33320

A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-638 <SER>

A;Cross-references: GB:M30282; NID:G205010; PIDN:AAA41463.1; PID:G205011

A;Note: part of this sequence, including the amino ends of both the heavy and light chain

R;Paquin, J.; Benjannet, S.; Sawyer, N.; Lazure, C.; Chretien, M.; Seidah, N.G.

Biochim. Biophys. Acta 999, 103-110, 1989

A;Title: Rat plasma kallikrein: purification, NH(2)-terminal sequencing and development

A;Reference number: S06851; MUID:90089457; PMID:2597701

A;Accession: S06851

A;Molecule type: Protein

A;Residues: 20-45;391-413 <PAQ>

R;Seidah, N.G.; Ladenheim, R.; Mbikay, M.; Hamelin, J.; Lutfalla, G.; Rougeon, R.; Lazur

DNA Cell Biol. 8, 563-574, 1989

A;Title: The cDNA structure of rat plasma kallikrein.

A;Reference number: I53041

A;Accession: I53041

A;Status: translated from GB/EMBL/DBJ

A;Residues: 1-638 <RES>

A;Cross-references: GB:M58590; NID:G206721; PIDN:AAA42069.1; PID:G206722

C;Comment: This protein, synthesized in the liver, circulates as a noncovalent complex w

C;Comment: The zymogen is activated by factor XIa, which cleaves the molecule into a li

are linked by one or more disulfide bonds.

C;Genetics:

A;Gene: PK

C;Superfamily: coagulation factor XI; trypsin homology

C;Keywords: blood coagulation; duplication; fibrinolysis; glycoprotein; hydrolase; infla

F;1-19/Domain: signal sequence #status predicted <SIG>

F;20-390/Product: plasma kallikrein heavy chain #status experimental <NATI>

F;110-199/Domain: apple repeat <AP1>

F;200-289/Domain: apple repeat <AP2>

F;291-380/Domain: apple repeat <AP3>

F;391-638/Product: plasma kallikrein light chain #status experimental <NAT2>

F;391-621/Domain: trypsin homology <TRY>

F;21-104/47-71;51-57;111-194;137-166;141-147;201-284;227-256;231-237;292-375;318-347;322

F;127;215;308;453;459;494/Binding site: carbohydrate (Asn) (covalent) #status predicted

F;396/Binding site: carbohydrate (Asn) (covalent) #status experimental

F;434;483;578/Active site: His, Asp, Ser #status predicted

Query Match 17.1%; Score 386.5; DB 1; Length 638;

Best Local Similarity 31.7%; Pred. No. 2.9e-22;

Matches 89; Conservative 58; Mismatches 109; Indels 25; Gaps 6

Qy 126 CMVHDCADGKLFQCGQKTLPRFKIIGTEFTTIENQWFAAIYRRHRRGGVYVCGGSL 18;

Db 375 CKVVESSD-----C---TTKINARIVGTTSSSLGEPWQVSL--QVXLVSNHMCQGS 42;

Qy 186 ISPCWVISATHCFFIDYPKKEDYIVYLGSRLSNNTQGBMKFEVENLIHKDYSADTLAHH 24;

Db 424 IGRQWILTAACFDGIPYDPVWRIYGGILNSETNTKTPFSSIKELIIHQYKMGSEGY- 48;

Qy 246 NDIALLKIRSKEGRCQAQPSRTIQTICLPSMYNDPQFGTSCBITGFG---KENSTDYLYPE 30;

Db 483 -DIALIKLQTP---LNVTEFKPKICLSKADNTIYTNWVTWGTGTYKERGETQNI--- 53;

Qy 303 QLKMTVVKLISHRECOQPHYVGVSEVTTQMLCAAPQWKTDCQDGGPLVCSLQGRWTL 36;

Db 535 -LQKATIFLVNFEBQCK-KYRDYVITKQMICAGYKEGIDACKGDSGGPLVCKHSGRWQL 59;

Qy 363 TGIVSWGRCGALKDQPGVYTVSHPLPWIRSHTRKEENGLAL 403

Db 593 VGITSWGEGCARKEQPGVYTVKVAEIDWILEKIQSSKERAL 633

RESULT 45

KXHU

protein C (activated) (EC 3.4.21.69) precursor - human

N;Alternate names: autoproteolytic protein C; plasma protein C

C;Species: Homo sapiens (man)

C;Date: 17-Mar-1987 #sequence revision 17-Mar-1987 #text\_change 16-Jul-1999

C;Accession: A22331; A25426; A21781; A23789; A00927

R;Foster, D.C.; Yoshitake, S.; Davie, E.W.

Proc. Natl. Acad. Sci. U.S.A. 82, 4673-4677, 1985

A;Title: The nucleotide sequence of the gene for human protein C.

A;Reference number: A22331; MUID:85270390; PMID:2991887

A;Accession: A22331

A;Molecule type: DNA

A;Residues: 1-461 <FOS1>

A;Cross-references: GB:M11228; NID:G190333; PIDN:AAA60166.1; PID:G190334

R;Plutsky, J.; Hoskins, J.A.; Long, G.L.; Crabtree, G.R.

Proc. Natl. Acad. Sci. U.S.A. 83, 546-550, 1986

A;Title: Evolution and organization of the human protein C gene.

A;Reference number: A25426; MUID:86120978; PMID:3511471

A;Accession: A25426

A;Molecule type: DNA

A;Residues: 1-445; 'L', 446-461 <PUJ>

A;Cross-references: GB:M12712; NID:G190330; PIDN:AAA60165.1; PID:G190332

R;Foster, D.; Davie, E.W.

Proc. Natl. Acad. Sci. U.S.A. 81, 4766-4770, 1984

A;Title: Characterization of a cDNA coding for human protein C.

A;Reference number: A21781; MUID:84272714; PMID:6589623

A;Accession: A21781

A;Molecule type: mRNA

A;Residues: 'Q', 107-461 <FOS2>

A;Cross-references: GB:K02059; NID:G190322; PIDN:AAA60164.1; PID:G190323

R;Beckmann, R.J.; Schmidt, R.J.; Santerre, R.F.; Plutsky, J.; Crabtree, G.R.;

Nucleic Acids Res. 13, 5233-5247, 1985

A;Title: The structure and evolution of a 461 amino acid human protein C precu

A;Reference number: A23789; MUID:85269639; PMID:2991859

A;Accession: A23789

A;Molecule type: mRNA

A;Residues: 1-461 <BEC>

A;Cross-references: GB:X02750; NID:G35689; PIDN:CAA36528.1; PID:G763120

R;Miletich, J.P.; Broze Jr., G.J.

J. Biol. Chem. 265, 11397-11404, 1990

A;Title: Beta protein C is not glycosylated at asparagine 329. The rate of tra

A;Reference number: A44605; MUID:90293094; PMID:1694179

A;Contents: annotation; carbohydrate binding sites; activation peptide

A;Note: the alpha form of protein C is glycosylated at Asn-329, and the beta f

R;Harris, R.J.; Ling, V.T.; Spellman, M.W.

J. Biol. Chem. 267, 5102-5107, 1992

A;Title: O-linked fucose is present in the first epidermal growth factor domain

F:63.107.200/Active site: His, Asp, Ser #status predicted

Query Match 17.3%; Score 390.5; DB 2; Length 247;  
Best Local Similarity 38.1%; Pred. No. 5e-23;  
Matches 93; Conservative 41; Mismatches 83; Indels 27; Gaps 9;

QY 150 KIIGGFTTIENQ-PWFAATVRRHGGVTVVCGSLSPCWVISAATHCFIDYPKKEDYI 208  
DB 23 KIVGG-YTCAENSVQVSLNAGY-----HFCGSLNDQWVSAACHY-----QYHIQ 70

QY 209 VYLGSRSLNSNTQGEMKFEVENLILHKDYSADTLAHHNDIALLKIRSKEGRCAPSRITQ 268  
DB 71 VELGEYNDVLEGGQFIDASKIIRPKYSSWTL--DNIDILIKLSTP-----AVINARVS 124

QY 269 TICLPSMNDPQFTSCITGFKENSTDYYPBOLKMTVVKLISHRECQOPHYGSEVT 328  
DB 125 TLLPSAC--ASAGTECLISGMNTLSSGVNYPDLQLCVAPLISHADCEAS--YPGQIT 180

QY 329 TQMLCAADPQWKTDSCQDSGGPLVCSLQGRMTLTGIVSGRGCAKDKPKGVYTRVSHFL 388  
DB 181 NNMICAGFLEGGKDCQDSGGPVACNQ-----LQIVSWGYGCAQKGPVYTKVCNYV 236

QY 389 PMIR 392  
DB 237 DWIQ 240

## RESULT 42

A21195  
Chymotrypsin (EC 3.4.21.1) 2 precursor - dog  
C:Species: Canis lupus familiaris (dog)  
C:Date: 27-Jul-1990 #sequence\_revision 27-Jul-1990 #text\_change 22-Jun-1999  
C:Accession: A21195  
R:Pinsky, S.D.; LaForge, K.S.; Luc, V.; Scheele, G.  
Proc. Natl. Acad. Sci. U.S.A. 80, 7486-7490, 1983  
A:Title: Identification of cDNA clones encoding secretory isoenzyme forms: sequence de  
A:Reference number: A21195; MUID:84170253; PMID:6584866  
A:Accession: A21195  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-263 <PIN>  
A:Cross-references: GB:K01173; NID:g163945; PIDN:AAA30841.1; PID:g163946  
A:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; protein digestion; serine proteinase  
F:34-256/Domain: trypsin homology <TRY>  
F:75,120,213/Active site: His, Asp, Ser #status predicted

Query Match 17.3%; Score 390; DB 2; Length 263;  
Best Local Similarity 33.2%; Pred. No. 5.8e-23;  
Matches 86; Conservative 51; Mismatches 100; Indels 22; Gaps 7;

QY 138 FCGGKTLRPF-----KIIGGFTTIENQPFPAATVRRHGGSVTVVCGSLSPCWVIS 193  
DB 17 FCGGVPAIQVLSGLSRVINGEDAVGSPWQVSL---QDSTGFHFCGSLISEDWVVT 72

QY 194 ATHCFIDYPKKEDYIVLGRSLNSNTGEMKFEVENLILHKDYSADTLAHHNDIALLKI 253  
DB 73 AAHCGV----RTHQVAGEFQGGDASIQVLKIAKVPKPKFNFTI--NNDITLKL 126

QY 254 RSKEGRCAQPSRIITICLPSMYNDPQFTSCITGFKENSTDYYPBOLKMTVVKLIS 313  
DB 127 ATP-----ARFSKTVSAVCLPQATDDFPAGTLCTVTGWLTKHTNANTPKLQQAALPLLS 182

QY 314 HRECQOPHYGSEVTIKMLCAADPQWKTDSCQDSGGPLVCSLQGRMTLTGIVSGRCA 373  
DB 183 NAECKK--FWGSKITDLMWCAGAS--GVSSCMGSDGGPLVCOQGAWTLVGIVSGSGTC 238

QY 374 LKDKPGVYTRVSHFLPWIR 392  
DB 239 STSTPGVYARVTKLIPWQ 257

## RESULT 43

F:63.107.200/Active site: His, Asp, Ser #status predicted

Query Match 17.2%; Score 387.5; DB 1; Length 407;  
Best Local Similarity 27.7%; Pred. No. 1.5e-22;  
Matches 117; Conservative 54; Mismatches 139; Indels 113; Gaps 20;

QY 13 CLNGGTCVSNKYPSNIHWCNCPKFGQHCHIDKSK--TCYEGNGHFYRGKASTDTMGRP 70  
DB 55 CQGGSC-EDQLRSYI--CFCPDGFEGRCNCTDKQSOLICANDNG----- 96

QY 71 CLPWSATVLTQYTHAHRSDALQLGLGKHNYC-RNPDRRRPWCYVQVGLKPLVQECMVH 129  
DB 97 -----GCEQYQADPGAGRCFVCHCEYALQ----- 121

QY 130 DCADG-----KLKFCQGKTL-----RPRKIIGGFTTIENQPFPAATVRRHGGSV 177  
DB 122 --ADGVSCAPTVEYPCGKIPVLEKRNKSPQGRIVSGHVCPKGECPWQAML---KLNGAL 176

QY 178 TVYCGSLSPCWVISAATHCFIDYPKKEDYIVLGRSLNSNTGEMKFEVENLILHKDY 237  
DB 177 --LCGGTGLGPAWVSSAAHCFELRLSRGNLTAVLGSHDLSRVGEPQERRVQAQIIVFKQY 234

QY 238 SADTLAHHNDIALLKIRSKEGRCAPQ-----SRITQITCLPSMYNDPQFTS-----CEI 287  
DB 235 VPGQTDH--DVALQL-----AQFVALGDHVAPLCLP---DPDFADQTLAFVRSFV 281

QY 288 TGFQK--ENSTDYLYPEQLKMTVVKLISHREC-----QOPHYGSEVTIKMLCAADPQWK 340  
DB 288 TGFQK--ENSTDYLYPEQLKMTVVKLISHREC-----QOPHYGSEVTIKMLCAADPQWK 340

QY 390 WIRSH 395  
Db 848 WIKQ 853

RESULT 39  
S55066  
trypsin (EC 3.4.21.4) II precursor, pancreatic (clone 2-P29) - chicken  
N:Alternate names: trypsinogen II  
C:Species: Gallus gallus (chicken)  
C>Date: 23-Aug-1995 #sequence\_revision 19-Oct-1995 #text\_change 22-Jun-1999  
C:Accession: S55066; S72347  
R:Wang, K.; Gan, L.; Lee, I.; Hood, L.  
Biochem. J. 307, 471-479, 1995  
A:Title: Isolation and characterization of the chicken trypsinogen gene family.  
A:Reference number: S55065; MUID:95251611; PMID:7733885  
A:Accession: S55066  
A:Molecule type: mRNA  
A:Residues: 1-248 <WAN1>  
A:Cross-references: EMBL:U15157; NID:G603906; PIDN:AAA79914.1; PID:G603907  
A:Experimental source: clone 2-P29  
A:Accession: S72347  
A:Molecule type: DNA  
A:Residues: 1-248 <WAN2>  
A:Cross-references: EMBL:U15157; NID:G603906; PIDN:AAA79914.1; PID:G603907  
A:Experimental source: clone 2-P29  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen  
F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-25/Domain: activation peptide #status predicted <APT>  
F:26-248/Product: trypsin II #status predicted <WAT>  
F:26-241/Domain: trypsin homology <TRY>  
F:65,109,202/Active site: His, Asp, Ser #status predicted

Query Match 17.5%; Score 395.5; DB 2; Length 248;  
Best Local Similarity 38.7%; Pred. No. 2e-23;  
Matches 94; Conservative 40; Mismatches 94; Indels 25; Gaps 7;

QY 150 KIIGFEFTIENQWFAALYRRHGGSVYVCGSLSPCWVISATHCFIDYPKEDYIV 209  
Db 25 KIVGYTCPEHSPVQVSL-----NSGYHFCGGLNSQWLVLSAAHCY-----KSRIOV 73

QY 210 YLGRSLNSNTQGMKFEVENILHKDYSADTLAHNDIALKLRSEKRCQAPSRITQT 269  
Db 74 RLGENIVDQDESEVRRSSVIRHPKYSITL--NNDIMLIKAS-----AVEVSADIQP 127

QY 270 ICLPSMYNDPQGTSCETGFGKENSTDYLYPEQLKMTVVKLISHRECCQPHYVSEVTT 329  
Db 128 IALPS--SCARAGTECLISGWNTLSNGYNYPELLQCLNAPILSDQECQEA--YPGDITS 183

QY 330 KMLCAADPQWKTDSCQDSGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTVRSHFLP 389  
Db 184 NMICVGFLEGGKDCSQDSGGPVCNGE----LQIVSWGIGCALKGYPGYVTKVCNYVD 239

QY 390 WIR 392  
Db 240 WIQ 242

RESULT 40  
TREOTR  
trypsin (EC 3.4.21.4) precursor - bovine  
N:Contains: trypsinogen  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 24-Apr-1984 #sequence\_revision 28-Feb-1986 #text\_change 18-Jul-1997  
C:Accession: A90164; A00946; S08774  
R:Mike, O.; Holeysovsky, V.; Tomasek, V.; Sorm, F.  
Biochem. Biophys. Res. Commun. 24, 346-352, 1966  
A:Title: Covalent structure of bovine trypsinogen. The position of the remaining amides.  
A:Reference number: A90164; MUID:67168848; PMID:5967094  
A:Accession: A90164  
A:Molecule type: protein  
A:Residues: 1-57, 'Q', 59-67, 'Q', 69-150, 'N', 152-176, 'N', 178-229 <MIK>

R:Hartley, B.S.  
Philos. Trans. R. Soc. Lond. B257, 77-87, 1970  
A:Reference number: A93755  
A:Contents: annotation; revisions  
R:Titani, K.; Ericsson, L.H.; Neurath, H.; Walsh, K.A.  
Biochemistry 14, 1358-1366, 1975  
A:Title: Amino acid sequence of dogfish trypsin.  
A:Reference number: A00950; MUID:75146445; PMID:1092332  
A:Contents: annotation; revisions that shown  
A:Note: the sequence agrees with that shown  
R:Bode, W.; Schwager, P.  
J. Mol. Biol. 98, 693-717, 1975  
A:Title: The refined crystal structure of bovine beta-trypsin at 1.8 angstrom  
A:Reference number: A92954; MUID:7607097; PMID:512  
A:Contents: annotation; X-ray crystallography; binding sites for calcium, sub  
ate, a te  
C:Comment: Trypsinogen is synthesized in the acinar cells of the pancreas.  
C:Comment: Autocatalytic cleavage after Lys-6 leads to beta-trypsin by releas  
s pseudotrypsin. A cleavage may also occur after Arg-105.  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen  
F:1-229/Product: trypsinogen #status experimental <2YM>  
F:1-6/Domain: activation peptide #status experimental <APT>  
F:7-222/Domain: trypsin homology <TRY>  
F:7-131,132-229/Product: alpha-trypsin #status experimental <MPT>  
F:6-7/Cleavage site: Lys-11e (enteropeptidase) #status experimental  
F:13-143,31-47,115-216,122-189,154-168,179-203/Disulfide bonds: #status experi  
ntal  
F:46,90,183/Active site: His, Asp, Ser #status experimental  
F:58,60,63,68/Binding site: calcium (Glu, Asn, Val, Glu) #status experimental  
F:131-132/Cleavage site: Lys-Ser (autolytic) #status experimental

Query Match 17.5%; Score 394.5; DB 1; Length 229;  
Best Local Similarity 37.4%; Pred. No. 2.2e-23;  
Matches 91; Conservative 41; Mismatches 86; Indels 25; Gaps 8;

QY 150 KIIGFEFTIENQWFAALYRRHGGSVYVCGSLSPCWVISATHCFIDYPKEDYIV 209  
Db 6 KIVGYTCGANTVPQVSL-----NSGYHFCGGLNSQWLVLSAAHCY-----KSGIOV 54

QY 210 YLGRSLNSNTQGMKFEVENILHKDYSADTLAHNDIALKLRSEKRCQAPSRITQT 269  
Db 55 RLGEDINIVVEGNEQFISASKSIVHPSYNSNTL--NNDIMLIKAS-----AASLNSRVAS 108

QY 270 ICLPSMYNDPQGTSCETGFGKENSTDYLYPEQLKMTVVKLISHRECCQPHYVSEVTT 329  
Db 109 ISLPT--SCASAGTQCLISGWNTLSNGYNYPELLQCLNAPILSDQECQEA--YPGDITS 164

QY 330 KMLCAADPQWKTDSCQDSGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTVRSHFLP 389  
Db 165 NMFCAGYLEGGKDCSQDSGGPVCNS--GK--LQIVSWGSGCAQKRGVYTKVCNYVS 220

QY 390 WIR 392  
Db 221 WIK 223

RESULT 41  
S13813  
trypsin (EC 3.4.21.4) - bovine  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 02-Dec-1993 #sequence\_revision 03-Aug-1995 #text\_change 22-Jun-1999  
C:Accession: S13813  
R:le Huerou, I.; Wicker, C.; Guilloreau, P.; Touleec, R.; Puigserver, A.  
Eur. J. Biochem. 193, 767-773, 1990  
A:Title: Isolation and nucleotide sequence of cDNA clone for bovine pancreatic  
A:Reference number: S13813; MUID:91065383; PMID:1701147  
A:Accession: S13813  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-247 <HUE>  
A:Cross-references: EMBL:X54703; NID:9829; PIDN:CAA38513.1; PID:9830  
C:Superfamily: trypsin; trypsin homology  
C:Keywords: hydrolase; protein digestion; serine proteinase  
F:24-239/Domain: trypsin homology <TRY>





A:Accession: A36557  
A:Molecule type: mRNA  
A:Residues: 1-638 <SEI>  
A:Cross-references: GB:M58588; NID:G200358; PIDN:AAA63393.1; PID:G200359  
A:Note: part of this sequence, including the amino ends of both the heavy and light chain  
C:Comment: This protein, synthesized in the liver, circulates as a noncovalent complex  
C:Comment: The zymogen is activated by factor XIIa, which cleaves the molecule into a light  
are linked by one or more disulfide bonds.  
C:Superfamily: coagulation factor XI; trypsin homology  
C:Keywords: blood coagulation; duplication; fibrinolysis; glycoprotein; hydrolase; infla  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-390/Product: plasma kallikrein heavy chain #status experimental <HCH>  
F:20-109/Domain: apple repeat <AP1>  
F:110-199/Domain: apple repeat <AP2>  
F:200-289/Domain: apple repeat <AP3>  
F:291-380/Domain: apple repeat <AP4>  
F:391-621/Product: plasma kallikrein light chain #status experimental <LCH>  
F:391-621/Domain: trypsin homology <TRY>  
F:21-104,47-77,51-57,111-194,137-166,141-147,201-284,227-256,231-237,292-375,318-347,322  
F:127,215,308,396,494/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:434,483,578/Active site: His, Asp, Ser #status predicted

Query Match 18.18; Score 407.5; DB 1; Length 638;  
Best Local Similarity 33.1%; Pred. No. 6.7e-24;  
Matches 93; Conservative 56; Mismatches 107; Indels 25; Gaps 8;

QY 126 CMWHDGDKLKFQCGOKTLRPRFKIIGGEFTTIENQPFALYRRHGGSVTVYVCGSL 185  
Db 375 CKLVSDPD-----C---TTKINARIVGGTNASLGEWPMQVSL--QVKLVSTQHLGGSI 423

QY 186 ISPCWVLSATHCFIDYFKKEDYIVYLGSRSLNNTQEMKFEVENILHKDYADTLAH 245  
Db 424 IGRQVLTAAHCFDIPYDPMWRIYGGILSLSEITKTPSSRIKELIIHOEYKVS--EGN 481

QY 246 NDIALKIKLQTP----LNYTFQRPICLPKADTIYTCWTVGWYTKQEGTQNI--- 534  
Db 482 YDIALIKLQTP----LNYTFQRPICLPKADTIYTCWTVGWYTKQEGTQNI--- 534

QY 303 QKMTVVKLISHRCQPHYGSVTTMCLCAADPQWTKDSQDGGPLVCSLQGRMTL 362  
Db 535 -LQKATPLPVNESQK-KYRDYVINKOMICAGYKEGGTDACKDGGSGPLVCKHSGRWQL 592

QY 363 TGIWSWGRGALKDKPGVYTRVSHFLPWIRSHTKENGLAL 403  
Db 593 VGIISWEGGCKRQDPGVYTKVSEYMDWILEKTQSSDVRL 633

RESULT 35  
S00845  
hepsin (EC 3.4.21.-) - human  
C:Species: Homo sapiens (man)  
C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 18-Jun-1999  
C:Accession: S00845  
R:Levytus, S.P.; Loeb, K.R.; Hagen, F.S.; Kurachi, K.; Davie, E.W.  
Biochemistry 27, 1067-1074, 1988  
A:Title: A novel trypsin-like serine protease (hepsin) with a putative transmembrane dom  
A:Reference number: S00845; MUID:98209431; PMID:2835076  
A:Accession: S00845  
A:Molecule type: mRNA  
A:Residues: 1-417 <LEY>  
A:Cross-references: EMBL:X07732; NID:G32063; PIDN:CAA30558.1; PID:G32064  
C:Genetics:  
A:Gene: GDB:HPN; TMPSR1; hepsin  
A:Cross-references: GDB:I35685; OMIM:142440  
A:Map position: 19q11-19q13.2  
C:Superfamily: hepsin; trypsin homology  
C:Keywords: hydrolase; liver; serine proteinase; transmembrane protein  
F:23-45/Domain: transmembrane #status predicted <TMN>  
F:163-400/Domain: trypsin homology <TRY>  
F:188-204,291-359,322-338,349-381/Disulfide bonds: #status predicted  
F:203,257,353/Active site: His, Asp, Ser #status predicted

Query Match 17.98; Score 404.5; DB 1; Length 417;

Best Local Similarity 34.9%; Pred. No. 7.2e-24;  
Matches 106; Conservative 55; Mismatches 98; Indels 45; Gaps 1;

QY 122 LVQECWHDGCDGK-----LKQCGOKTLRPRFKIIGGEFTTIENQPFALYRRHGGSV 177  
Db 131 LLEVISVDCPRGRFLAALQCDCGRRLK-PVDRIVGGEDTSLGRWPMQVSL--RYDG--- 184

QY 178 TVVCGSLISPCWVLSATHCFIDYFKKEDYIVYLGSRSLNNTQGE-----MKFEVENLI 232  
Db 185 AHLCCGSLSDGMDWLTAAHCF---PERNR---VLSRWVFAGAAQASPHGLQLGQVAV 238

QY 233 LHKDY-----SADTLAHNDIALKIRSEKGRCAQPSRTIOTICLPMSYNDPQFGTSCBIT 288  
Db 239 YHGGVLPFRDPNSSENSNDIALVHLSPP-----LPLTEIYQVCLPAAGQALVDGKICVT 294

QY 289 GFKENSTDIYYPEQ-----LKMVTVKLVISHRCQPHYGSVTTMCLCAADPQWTKDSQC 345  
Db 295 GNG---NTQY-YGQAGVLOARVPIISNDVNCNGADFYGNQIKPMFCAGYPEGIGIDACQ 350

QY 346 GDSGGPLVC---SLQGRMTLTGVSWGRGALKDKPGVYTRVSHFLPW-----IRSHTKS 397  
Db 351 GDSGGPFVCEDSISRTPRWRLCGIVSWGTGALCAQKPGVYTKVSDFRFWIFQALIKHS-E 409

QY 398 ENGL 401  
Db 410 ASGM 413

## RESULT 36

## KOHUP

plasma kallikrein (EC 3.4.21.34) precursor - human  
N:Alternate names: kininogenin; plasma prekallikrein  
C:Species: Homo sapiens (man)  
C:Date: 13-Aug-1986 #sequence\_revision 13-Aug-1986 #text\_change 18-Jun-1999  
C:Accession: A00921; A37939  
R:Chung, D.W.; Fujikawa, K.; McMullen, B.A.; Davie, E.W.  
Biochemistry 25, 2410-2417, 1986  
A:Title: Human plasma prekallikrein, a zymogen to a serine protease that conta  
A:Reference number: A00921; MUID:86243359; PMID:3521732  
A:Accession: A00921  
A:Molecule type: mRNA  
A:Residues: 1-638 <CHU>  
A:Cross-references: GB:M13143; NID:G190262; PIDN:AAA60153.1; PID:G190263  
R:McMullen, B.A.; Fujikawa, K.; Davie, E.W.  
Biochemistry 30, 2050-2056, 1991  
A:Title: Location of the disulfide bonds in human plasma prekallikrein: the p  
A:Reference number: A37939; MUID:91152016; PMID:1998666  
A:Accession: A37939  
A:Molecule type: protein  
A:Residues: 20-27,40-46, 'X', 48, 'H', 50, 'X', 52-70, 'H', 75-76, 'X', 78-80,103-113,113  
:480-283, 'X', 285;287-291, 'X', 293-295;314-317, 'X', 319-320;321-324, 'X', 329-333;3  
525;538-551;562, 'X', 564-567;573, 'X', 575-576;578-583, 'X', 585;592-604 <MCW>  
C:Comment: This protein, synthesized in the liver, circulates as a noncovalent  
C:Comment: The zymogen is activated by factor XIIa, which cleaves the molecule  
are linked by one or more disulfide bonds.  
C:Comment: The enzyme cleaves Lys-Arg and Arg-Ser bonds. It activates, in a re  
inogen and may also play a role in the renin-angiotensin system by converting  
C:Genetics:  
A:Gene: GDB:KLK3  
A:Cross-references: GDB:I27575; OMIM:229000  
A:Map position: 4q35-q35  
C:Superfamily: coagulation factor XI; trypsin homology  
C:Keywords: blood coagulation; duplication; fibrinolysis; glycoprotein; hydrol  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-638/Product: plasma kallikrein heavy chain #status predicted <HCH>  
F:20-390/Domain: plasma kallikrein heavy chain #status predicted <MAT>  
F:20-109/Domain: apple repeat <AP1>  
F:110-199/Domain: apple repeat <AP2>  
F:200-289/Domain: apple repeat <AP3>  
F:291-380/Domain: apple repeat <AP4>  
F:391-638/Domain: plasma kallikrein light chain #status predicted <LCH>  
F:391-621/Domain: trypsin homology <TRY>  
F:21-104,47-77,51-57,111-194,137-166,141-147,201-284,227-256,231-237,292-375,3

QY 265 RTIOTICLPNMYNDPO-FGTSCHITGKFNSTDIYLPQKMTVVVKLISHRECOQPHY 323  
Db 630 THVLPACLPLWRERPOKTSNCHITGWG---DTGRAYSRTLQAAVPLLPKRFCKE--SY 684  
QY 324 GSEVTTWMLCAADPOW--KTDSCGSGPLVCSLQGR-WTLTGIVSWGEGCALDKPKGV 380  
Db 685 KGLFTGMLCAGNLQENRVDSCGDSGGPLMCKEPDESVMVYTGVTSMGVCGRKDTPGV 744  
QY 381 YTRVSHPLPWIRSHST 395  
Db 745 YTRVPAFVPMIKSVT 759

RESULT 32  
146260  
plasmin (EC 3.4.21.7) precursor - western European hedgehog  
C/Species: Erinaceus europaeus (western European hedgehog)  
C/Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 16-Jul-1999  
C/Accession: I46260  
J/Lawn, R.M.; Boommark, N.W.; Schwartz, K.; Lindahl, G.E.; Wade, D.P.; Byrne, C.D.; Fong  
J. Biol. Chem. 270, 24004-24009, 1995  
A/Title: The recurring evolution of Lp(a): Insights from cloning of hedgehog apolipoprotein  
A/Reference number: I46259; MUID:96025778; PMID:7592597  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-810 <LAW>  
A/Cross-references: EMBL:U33171; NID:gl046360; PID:gl046361  
C/Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology  
C/Keywords: hydrolyase; serine proteinase  
F:1-96/Domain: plasminogen-related protein precursor homology <PLPR>  
F:103-181/Domain: kringle homology <KR1>  
F:185-262/Domain: kringle homology <KR2>  
F:275-352/Domain: kringle homology <KR3>  
F:373-456/Domain: kringle homology <KR4>  
F:482-561/Domain: kringle homology <KR5>  
F:582-803/Domain: trypsin homology <TRY>

Query Match 19.1%; Score 431; DB 2; Length 810;  
Best Local Similarity 30.7%; Pred. No. 1.3e-25;  
Matches 122; Conservative 49; Mismatches 157; Indels 70; Gaps 16;

QY 27 NIHW--CNCPKFGGQCEIDKS-----KTCYENGHFFYRGKASTDTMRPCL 72  
Db 445 SVRWEFCNLKCKSGTMSATNSPQVSSASESEQDCIIDNGKXRGTRKATTGAGTPCQ 504  
QY 73 FWNATVLQOTYHAH-----RSDALQLGLKHNVCNPD--NRERPWCYQVGLKPL 122  
Db 505 AWA-----QPHHSIFTPETNPAVL-----QENVCNPDGANGPNCYT--TNPRL 552  
QY 123 VOECMVHDCAGKLFQCGQKTLRPFKI-----IGGETTTENQWFAIYRRHRGGSVTY 179  
Db 553 FDYCDIPHCVSFS--SADCGKPKVEFK-KCPRGVGCVAHPHSWPQVSLRRFQ-----H 605  
QY 180 VCGGSLSPGWISATHCFIDYPKKEDIVYLG--SRSLNSNTQ--GEMKFEVENILH 234  
Db 606 FCGGLTLPSEWVTAHLCLEFSPAIYKVVLGAHQETRLERDVQIKGVTKMFL----- 660  
QY 235 KDYSADTLAHNDIALLKIRSGEACQPSRTIOTICLPNMYNDPOFGTSCBITGFOKEN 294  
Db 661 -----PYRADIALKLSSP-----AIIIDKHPACLPNSNVMVADRSLCYITGWGETK 708  
QY 295 STDYLYPQLKMTVVKLISHRECOQPHYVGSSEVTTKMLCAADPOWKTDSCGDSGGPLVC 354  
Db 709 GT--YGAGLLKEAQLPVFIENKQSNQSFNLRVSTELCAGHLAGGVDSGCGGGPLVC 766  
QY 355 SLQGRMTLTGIVSWGRCALDKPKGVYTRVSHFLPWIR 392  
Db 767 PEKDRVILQGVTSWGLGCARLTRPGVYVVRVSRVSWLQ 804

RESULT 33  
A57014

proctasin (EC 3.4.21.-) precursor - human  
C/Species: Homo sapiens (man)  
C/Date: 24-May-1996 #sequence\_revision 24-May-1996 #text\_change 21-Apr-2003  
C/Accession: A57014; A54866  
R/Yu, J.X.; Chao, L.; Chao, J.  
J. Biol. Chem. 270, 13483-13489, 1995  
A/Title: Molecular cloning, tissue-specific expression, and cellular localization of hu  
A/Reference number: A57014; MUID:95286644; PMID:7768952  
A/Accession: A57014  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-343 <RES>  
A/Cross-references: GB:L41351; NID:g862304; PIDN:ARC41759.1; PID:g862305  
A/Experimental source: prostate  
A/Note: parts of this sequence were determined by protein sequencing  
R/Yu, J.X.; Chao, L.; Chao, J.  
J. Biol. Chem. 269, 18843-18848, 1994  
A/Title: Proctasin is a novel human serine proteinase from seminal fluid. Purifi  
A/Reference number: A54866; MUID:94308140; PMID:8034638  
A/Accession: A54866  
A/Molecule type: protein  
A/Residues: 45-64 <YUA>  
C/Genetics:  
A/Gene: GDB:PRSS8  
A/Cross-references: GDB:676446; OMIM:600823  
A/Map position: 16p11.2-16p11.2  
C/Superfamily: trypsin; trypsin homology  
C/Keywords: glycoprotein; hydrolase; serine proteinase; transmembrane protein  
F:1-32/Domain: signal sequence #status predicted <SIG>  
F:33-44/Domain: product: proctasin #status predicted <MAT>  
F:45-281/Domain: proctasin light chain #status predicted <CHL>  
F:45-281/Domain: proctasin heavy chain #status predicted <CHH>  
F:323-341/Domain: trypsin homology <TRY>  
F:323-341/Domain: transmembrane #status predicted <TM1>  
F:327-354, 70-86, 168-244, 201-223, 234-262/Digulfide bonds: #status predicted  
F:85,134,236/Active site: His, Asp, Ser #status predicted  
F:159/Binding site: carbonylate (Asn) (covalent) #status experimental

Query Match 18.1%; Score 408.5; DB 1; Length 343;  
Best Local Similarity 36.6%; Pred. No. 2.8e-24;  
Matches 100; Conservative 36; Mismatches 100; Indels 37; Gaps 8;

QY 140 CQCKTLRFRFKITIGGETTTENQWFAIYRRHRGGSVTY-----VCGGSLSPGWISAT 195  
Db 37 CG--VAPQARTGSSAVAGQWPMQV-----SITYEGVHVCAGSLVSEQWLSAA 84  
QY 196 HCFIDYPKKEDIVYLGSRSLNSNTQGEWKFEVENILHKDYSADTLAHNDIALLKIRS 255  
Db 85 HCFPSEHHEKEAVEKLGHAQLDSYSEDAKVTLDKDIIPHPSYLOB--GSQGDIALQLQ-- 140  
QY 256 KEGRCAQPSRTIOTICLPNMYNDPOFGTSCBITGEGK--ENSTDYLYPQLKMTVVKLISH 314  
Db 141 --SRPITTSRYIRPICLPANASFPNGLHCTVTGHWGVAHSVSLTTPKPLQOLEVPLISR 198  
QY 315 REC-----QOPHYVGSSEVTTKMLCAADPOWKTDSCGDSGGPLVCISLQGRMTLTG 364  
Db 199 ETCNCLYINIDAKPESEPHF-----VQEDWVCAGVVEGKDACCGDSGGPLSCPVEGLWYLTG 254  
QY 365 IVSWGRCALDKPKGVYTRVSHFLPWIRSHYKE 397  
Db 255 IVSWGDAACGARNRPGVYTLASSYASWISQSKVTE 287

RESULT 34  
KQMSPL  
plasma kallikrein (EC 3.4.21.34) precursor - mouse  
C/Species: Mus musculus (house mouse)  
C/Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 18-Jun-1999  
C/Accession: A36557  
R/Seidah, N.G.; Sawyer, N.; Hamelin, J.; Mion, P.; Beaubien, G.; Brachpapa, L.;  
DNA Cell Biol. 9, 737-749, 1990  
A/Title: Mouse plasma kallikrein: cDNA structure, enzyme characterization, and  
A/Reference number: A36557; MUID:91090844; PMID:2264928

QY	50	CYEGNGHFYRGKASTDTMGPCPLPWNASATVLO--QTYHAHRSDALGLGKHNVCNPDN	107
Db	1068	CYHNGOSYRGTFSTTVTGHTCOSWSMTPHQKRTPENHPNDLTM-----NYCNPDA	1122
QY	108	RRRPWCYVQGLKPLVOE---CMVHDCAD-----GKLFQCCGQ	142
Db	1123	DTGWCFT--MDPSVRREYCNLTRCSDTEGTVTPTVPVPSLEAPSEQASSPDCGK	1179
QY	143	KTLRPR---FKLTGGFTTIENTOPFAAIRRHRRGGSVTVCGGSLIPSCWISATHCFI	199
Db	1180	PQVEPKKCPGSIYGGCVAHPSHPWQSL--RTRFGK--HFCGGLTSPENWITAAACLE	1235
QY	200	DYPKEDYIYVLGRSR---LNSNTQGBMKFEVENLIHKDYSADTLAHNDIALKIRSK	256
Db	1236	TFSRSPFYKILGAHQEWNLESHVQ---EIEVSRLFLEPGA-----DIALKL---	1281
QY	257	EGRCAQPSRTIQTICLPMSWNDQF-----GTSCEITGFGKENSTDYLYPQLKMTVVVKLI	312
Db	1282	-SRPAITDKVIPACLPS-----PNYITAMTECYITGWGETQTG--FGAGLLKEAQLHVI	1334
QY	313	SHRECOQPHYGVSEVTTKMLCAADPOWKTDSCGGDGGPLVCSIQGRWTLTGTVSWGRCG	372
Db	1335	ENIVCNHYEFLNGRVKSTELCAGHLAGGTDRCOQDNGGPPVCFDKKYLIRGITSWGPCC	1394
QY	373	ALKDKPGVYTRVSHFLPWI	391
Db	1395	ACPKNKPGVYVRVSFFVTWI	1413

RESULT 30

A:Accession: A61545  
A:Plasmin (EC 3.4.21.7) precursor - horse (fragments)  
N:Alternate names: plasminogen  
N:Contains: miniplasminogen  
C:Species: Equus caballus (domestic horse)  
C:Date: 28-Oct-1994 #sequence\_revision 01-Nov-1996 #text\_change 18-Jul-1997  
C:Accession: A61545; S17527  
R:Schaller, J.; Rickli, E.E.  
Enzyme 40, 63-69, 1988  
A:Title: Structural aspects of the plasminogen of various species.  
A:Reference number: A61545; MUID:89005015; PMID:3168975  
A:Accession: A61545  
A:Molecule type: protein  
A:Residues: 1-35/34-117 <SCH>  
R:Schaller, J.; Straub, C.; Kaempfer, U.; Rickli, E.E.  
Protein Seq. Data Anal. 4, 69-74, 1991  
A:Title: Complete amino acid sequence of equine miniplasminogen.  
A:Reference number: S17527; MUID:92052077; PMID:1946332

Query Match	20.08;	Score	451.5;	DB	2;	Length	455;
Best Local Similarity	32.5%;	Pred. No.	1.8e-27;				
Matches	122;	Conservative	45;	Mismatches	169;	Indels	39;
Gaps	12;						

  

Qy	27	NIRWCNPKFGGGHCIEDKSKTCYEGNHFYRGKASTDTMGRPCLPWNSATVLQOYYHA	86
Db	103	SVRWEFCNLRKCKSETVQEPSEPDCMLGIGYGGKATVTGRCQAWAA-----QEPHR	157
Qy	87	H----RSPALQGLGKHNYCNRPD--NRRRPVCYVQGLKPLVQECMVHDCADGKLKFCQG	142

```

Db      158 HSIFTEAPNWNLEKNYCRNPDGDNVGPWCYT--NPKQLFYCDVPOCESS--PFDGCK 214
Qy      143 KTLRPR--FKIIGGEFTIENQPFPAIYRHRGGSVTVYCGSLISPCWVISATHCFI 199
Db      215 PKVEPKCKSGRIVGGCAIAHSPW--QISLRTFG--RHFCGGTLISPEWVLTAAHGLE 270
Qy      200 DYPKEDYIVYLGSRSLNSNTGEMKFEVENLILHKDYSADTLAHNDIALLKIRSKGR 259
Db      271 RSRPSTYKVVLTGHLEPLAGAQIDVSKLFLPSRA-----DIALKLSS---- 318
Qy      260 CQAPSTIQTIT---CLPSSMYNDPQFGTSCITRFGKFNSTDYLYPEQLKNTVYKVLISHRE 316
Db      319 ---PAITQNVIPACLPADYVYVWAWBECFVTGWGE--TQDSSNAGVLKEAQLPIYENKV 373
Qy      317 CQPHYVYGBVTKMLCAADPOWKTDSCQDGGGLVCSLQGRMTLTGIVSGRCAUKD 376
Db      374 CNRYEYLVGRVSKTLCAGHLVGGVDSQDGGGLVCFEKDKYTLQGVTSWGLGCARP 433
Qy      377 KGVYTRVSHFLPWI 391
Db      434 KPGVYVRVSSFINWI 448

RESULT 31
JC5759
brain-specific serine proteinase (EC 3.4.21.-) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 24-Jan-1998 #sequence_revision 13-Mar-1998 #text_change 31-Mar-2000
C:Accession: JC5759
R:Yamamura, Y.; Yamashiro, K.; Tsuruoka, N.; Nakazato, H.; Tsujimura, A.; Yamaguchi, I.; Biochem. Biophys. Res. Commun. 239, 386-392, 1997
A:Title: Molecular cloning of a novel brain-specific serine protease with a kringle-1-like domain
A:Reference number: JC5759; MUID:98008848; PMID:9344839
A:Accession: JC5759
A:Molecule type: mRNA
A:Residues: 1-761 <YAM>
A:Cross-references: DDBJ:D89871
A:Experimental source: brain
C:Superfamily: brain-specific serine proteinase; scavenger receptor cysteine-rich domain
C:Keywords: glycoprotein; hydrolase; serine proteinase
F:85-157/Domain: kringle-like #status predicted <KRI>
F:163-266/Domain: scavenger receptor cysteine-rich domain homology <SRC7>
F:166-266, 273-372, 386-486/Domain: scavenger receptor cysteine-rich domain homology <SRC7>
F:513-516/Domain: furin binding #status predicted <FRB>
F:517-755/Domain: trypsin homology <TRY>
F:93,521,569/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:562,612,711/Active site: His, Asp, Ser #status predicted

```

RESULT 31

JC5759  
brain-specific serine proteinase (EC 3.4.21.-) - mouse  
C/Species: Mus musculus (house mouse)  
C/Date: 24-Jan-1998 #sequence\_revision 13-Mar-1998 #text\_change 31-Mar-2000  
C/Accession: JC5759  
R/Yamamura, Y.; Yamashiro, K.; Tsuruoka, N.; Nakazato, H.; Tsujimura, A.; Yamachi, I.  
Biochem. Biophys. Res. Commun. 239, 386-392, 1997  
A/Title: Molecular cloning of a novel brain-specific serine protease with a kringle-1 domain  
A/Reference number: JC5759; MUID:98008848; PMID:3344839  
A/Accession: JC5759  
A/Molecule type: mRNA  
A/Residues: 1-761 <YAM>  
A/Cross-references: DDBJ:D89871  
A/Experimental source: brain  
C/Suprafamily: brain-specific serine proteinase; scavenger receptor cysteine-rich domain  
C/Keywords: glycoprotein; hydrolase; serine proteinase  
F/85-157/Domain: kringle-like #status predicted <KRI>  
F/163-266/Domain: scavenger receptor cysteine-rich domain homology <SRC7>  
F/166-266, 273-372, 386-486/Domain: scavenger receptor cysteine-rich domain homology <SRC7>  
F/513-516/Domain: tyrpin binding #status predicted <FRB>  
F/517-755/Domain: tyrpin homology <TRY>  
F/93,521,569/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F/562,612,711/Active site: His, Asp, Ser #status predicted

Query Match	19.2%	Score	434	DB 2	Length	761			
Best Local Similarity	29.0%	Pred. No.	7.1e-26						
Matches	126	Conservative	62	Mismatches	159	Indels	88	Gaps	16
Qy	28	IHWNCNPKFGGQH-C--BIDSKTCY-EGNGH-----	56						
Db	346	VSFIQSRRQWGRHDCSHREDVGLTCYDPDSDGHRLLSPGPIRLVDGNNKKEGRVEVFVNG	405						
Qy	57	-----FYGKASTDTM---GRPCLPWNSATVLQTYHAHRSDA	91						
Db	406	QWGTICDDGWTDKAAVTCRQYGYKGPARTMAYFGGKGPIHMDNVKTCGTNEKALADC	465						
Qy	92	LQLGLGHNYCRPNDRRFPWCYVQ-----VGLKPLVQECMVHDCADGKLAFQCGQKTLR	146						
Db	466	VKQDIGRHN-CRHSFDAGVICDYLEKASSSGNKM-----LSSGCGRLRLH	511						
Qy	147	PREK-IIGCEFTTINQPFALYYRRHGGSVYYCGGSLISPCWVVSATHCFFIDY-PK	204						
Db	512	RQKRILIGNSLURGANPQASLRSRSHAGDGRLLCGATLLUSSCWLTAAHCFKRYGNNS	571						
Qy	205	EDYVYVGLSRLSNNTQGEMKFEVENLILHKDYSADTLAHHNDIALLKIRSKEGSCAOPS	264						
Db	572	RSVAVRGVYHTLVPEEFBOETIGVOOIVHRYNRPRDSY---DIALVQLQGPGECARLS	629						

A:Reference number: A47277; MUID:93165698; PMID:7679504  
A:Accession: A47277  
A:Status: preliminary; translation not shown; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-16 <RES>  
A:Cross-references: GB:107899; NID:G967973; PID:G967974  
R:Malgarutti, N.; Acquati, F.; Magnaghi, P.; Bruno, L.; Pontoglio, M.; Rocchi, M.; Sacco  
Proc. Natl. Acad. Sci. U.S.A. 89, 11584-11588, 1992  
A:Title: Characterization by yeast artificial chromosome cloning of the linked apolipoprotein  
A:Reference number: A47233; MUID:93087573; PMID:1454851  
A:Accession: 160956  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-16 <RES>  
A:Cross-references: GB:M90078; NID:G178786; PIDN:AAA35547.1; PID:G553188  
A:Note: apo(a) Gene 1 (nomenclature of reference I52415)  
A:Accession: A47233  
A:Status: preliminary; translation not shown; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-16 <RES>  
A:Cross-references: GB:M90079; NID:G178784; PIDN:AAA35546.1; PID:G553187  
R:Ichinose, A.  
Biochemistry 31, 3113-3118, 1992  
A:Title: Multiple members of the plasminogen-apolipoprotein(a) gene family associated with  
A:Reference number: 152415; MUID:92207924; PMID:1554698  
A:Accession: 152415  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-16 <RES>  
A:Cross-references: GB:M86877; NID:G178780; PIDN:AA849909.1; PID:G553185  
A:Note: apo(a) Gene 1 (nomenclature of reference I52415)  
A:Accession: 165286  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-16 <RES>  
A:Cross-references: GB:M86878; NID:G178782; PIDN:AAA51749.1; PID:G553186  
C:Genetics:  
A:Gene: GDB:152200  
A:Cross-references: GDB:120699; OMIM:152200  
A:Map position: 626-eq27  
A:Note: several genes closely linked on chromosome 6 are identical in the first coding exons of kringle repeats  
C:Superfamily: apolipoprotein(a); kringle homology; trypsin homology  
C:Keywords: hydrolase; kringle; lipid binding; lipoprotein; serine proteinase  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-4549/Product: apolipoprotein(a) #status experimental <MAT>  
F:28-105/Domain: kringle homology <KR1>  
F:142-219/Domain: kringle homology <KR2>  
F:256-333/Domain: kringle homology <KR3>  
F:370-447/Domain: kringle homology <KR4>  
F:484-561/Domain: kringle homology <KR5>  
F:598-675/Domain: kringle homology <KR6>  
F:712-789/Domain: kringle homology <KR7>  
F:826-903/Domain: kringle homology <KR8>  
F:940-1017/Domain: kringle homology <KR9>  
F:1054-1131/Domain: kringle homology <KR10>  
F:1168-1245/Domain: kringle homology <KR11>  
F:1282-1359/Domain: kringle homology <KR12>  
F:1396-1473/Domain: kringle homology <KR13>  
F:1510-1587/Domain: kringle homology <KR14>  
F:1624-1701/Domain: kringle homology <KR15>  
F:1738-1815/Domain: kringle homology <KR16>  
F:1852-1929/Domain: kringle homology <KR17>  
F:1966-2043/Domain: kringle homology <KR18>  
F:2080-2157/Domain: kringle homology <KR19>  
F:2194-2271/Domain: kringle homology <KR20>  
F:2308-2385/Domain: kringle homology <KR21>  
F:2422-2499/Domain: kringle homology <KR22>  
F:2536-2613/Domain: kringle homology <KR23>  
F:2650-2727/Domain: kringle homology <KR24>  
F:2764-2841/Domain: kringle homology <KR25>  
F:2878-2955/Domain: kringle homology <KR26>  
F:2992-3069/Domain: kringle homology <KR27>

F:3106-3183/Domain: kringle homology <KR28>  
F:3220-3297/Domain: kringle homology <KR29>  
F:3334-3411/Domain: kringle homology <KR30>  
F:3448-3525/Domain: kringle homology <KR31>  
F:3562-3639/Domain: kringle homology <KR32>  
F:3676-3753/Domain: kringle homology <KR33>  
F:3782-3859/Domain: kringle homology <KR34>  
F:3896-3973/Domain: kringle homology <KR35>  
F:4010-4087/Domain: kringle homology <KR36>  
F:4124-4201/Domain: kringle homology <KR37>  
F:4228-4307/Domain: kringle homology <KR38>  
F:4328-4541/Domain: trypsin homology <TRY>  
Query Match 21.2%; Score 478; DB 1; Length 4548;  
Best Local Similarity 36.0%; Pred. No. 1.9e-28;  
Matches 129; Conservative 36; Mismatches 133; Indels 60; Gaps 17;  
QY 50 CYEENGHVRGKASDTMTGRECLPNSATVLQYVYHAHRSALQLGLCK-----HNYCRN 104  
DB 4228 CMFGNGKGYRGKATVTVTGTCQEW-----AAQEPHRSHTFIPGINKWAGLEKNYCRN 4280  
QY 105 PD-NRRRPMCVYVGLKELVQECMVHDCADGKLKFOGQKTLRPR---FKIIGGEFTTIE 160  
DB 4281 PDGDIINGPWCYT-MNPRKLDYCDIPLCASS--SFDCKGPQVEPKKCPGSIIVGCVAPH 4337  
QY 161 NOPWFAAIYRRHGGSVTVVCGSLISPCWVISAHCIFIDYPKKEDYIVYLGSR---LN 217  
DB 4338 SWPQVSL--RTRFGK--HFCGTLISPEWVITAAHCKKSRPSSYKVIIGAHQEVNLE 4393  
QY 218 SNTQEMFEVENLILHRDYSADTLAHNDIALIKRSKEGRCAPSRITITICLPSPMYN 277  
DB 4394 SHVQ---EIEVSRLEFLEPTQA-----DIALKL-----SRPAVITDKVMPACLPS--- 4435  
QY 278 DPQF----GTSCEITGPKENSTDYLYPEQLKNTVVKLISHRECOQPHYGVSEVTTKMLC 333  
DB 4436 -PDYVMTARTCYITGNGETQGT--FGTGLLKBAQLLVIEVCN--HY-----KYIC 4483  
QY 334 AADPQWKTDSCGDSGGPLVCSLQGRMTLTGIVSGRCALKDKPGVYTRVSHFLPWI 391  
DB 4484 AEHLARGTDSGDSGGPLVCFKDKYILQGVTSWGLSCARENKPGVYARVSFRFTWI 4541  
RESULT 29  
A32869  
apolipoprotein(a) (EC 3.4.21.1) - rhesus macaque (fragment)  
C:Species: Macaca mulatta (rhesus macaque)  
C:Date: 22-Nov-1989 #sequence\_revision 22-Nov-1989 #text\_change 22-Jun-1999  
C:Accession: A32869; A30848  
R:Tomlinson, J.E.; McLean, J.W.; Lawn, R.M.  
J. Biol. Chem. 264, 5957-5965, 1989  
A:Title: Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of synthesis.  
A:Reference number: A32869; MUID:89174660; PMID:2925643  
A:Accession: A32869  
A:Molecule type: mRNA  
A:Residues: 1-1420 <TOM>  
A:Cross-references: GB:J04635; NID:G342072; PIDN:AAA36833.1; PID:G342073  
C:Superfamily: apolipoprotein(a); kringle homology; trypsin homology  
C:Keywords: hydrolase; kringle; lipid binding; lipoprotein; serine proteinase  
F:50-127/Domain: kringle homology <KR1>  
F:164-241/Domain: kringle homology <KR2>  
F:278-355/Domain: kringle homology <KR3>  
F:392-469/Domain: kringle homology <KR4>  
F:506-583/Domain: kringle homology <KR5>  
F:620-697/Domain: kringle homology <KR6>  
F:726-803/Domain: kringle homology <KR7>  
F:840-917/Domain: kringle homology <KR8>  
F:954-1031/Domain: kringle homology <KR9>  
F:1068-1145/Domain: kringle homology <KR10>  
F:1191-1413/Domain: trypsin homology <TRY>

Query Match 21.1%; Score 477; DB 2; Length 1420;  
Best Local Similarity 34.0%; Pred. No. 6.4e-29;  
Matches 129; Conservative 43; Mismatches 137; Indels 70; Gaps 15;

QY 57 FYFGKASTDTMGPRCLPWNATVLQOYTHAH-----RSDALQGLGKNNYCRNPD- 106  
Db 488 DYRGKTAATAAGTFCQGWAA-----QEPHRSIFTPQTNPRADL-----ENKYNCRNPDG 536  
QY 107 NRARPWCYVQVGLKPLVQECMHDCADGKLKFCQCKOTLRPR---FKIIGGEFTTIENOP 163  
Db 537 DVNGPWCYT-TNPKLYDYCIDPLCAS-SFGCKPQVEPKPCQGVVGGCVANPHSWP 594  
QY 164 WFAIYRRHRGGSVTVVCGSLSPCWISATHCFIDYPKEDYIYVLRSLNSNTQGE 223  
Db 595 WQISLRTFTG---QHFCGGTLAPBWLTAAHCKLSRPPFYKVLGAHEEYIRGLDV 651  
QY 224 MKEVENLILHKDYSADTLAHNDIALKIRSKGRCAPSRITQICLPSMYNDPOF-- 281  
Db 652 QEISVAKLLE-----PNNRDIALKL-----SRPATITDKVIPACLPSS-----PNYMV 695  
QY 282 --GTSCEITGFGKENSTDYLPQKLVTVVVKLISHRECQPHYGVGSEVTTKMLCAADPQW 339  
Db 696 ADRTICVITGWTGT--FGAGRLKEAQLPVLENKVRVEYLNVRKSTELCAGQLAG 753  
QY 340 KTDSCQDGGPLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWI 391  
Db 754 GVDSCQDGGPLVCFEKDKYILQGVTSWGLGCRPNKPGVYVRSRFDWI 805

## RESULT 27

## PLBO

Plasmin (EC 3.4.21.7) precursor - bovine  
N/Alternate names: plasminogen  
C/Species: Bos primigenius taurus (cattle)  
C/Date: 30-Sep-1997 #sequence revision 28-Apr-1995 #text\_change 18-Jun-1999  
C/Accession: S45046; A25835; I45961; S03736  
R/Berglund, L.; Andersen, M.D.; Petersen, T.E.  
submitted to the EMBL Data Library, May 1994  
A/Description: Cloning and characterization of the bovine plasminogen cDNA.  
A/Reference number: S45046  
A/Accession: S45046  
A/Molecule type: mRNA  
A/Residues: 1-812 <BER>  
A/Cross-references: EMBL:X79402; NID:G494962; PIDN:CAA55939.1; PID:G494963  
A/Experimental source: liver  
A/Note: it is uncertain whether Met-1 or Met-8 is the initiator  
R/Schaller, J.; Moser, P.W.; Dannegger-Muller, G.A.K.; Rossiet, S.J.; Kampfer, U.; Rick  
Eur. J. Biochem. 149, 267-278, 1985  
A/Title: Complete amino acid sequence of bovine plasminogen. Comparison with human plasmin  
A/Reference number: A25835; MUID:85203906; PMID:3846532  
A/Accession: A25835  
A/Molecule type: protein  
A/Residues: 27-334, D', 336-515, 'H', 517-554, 'L', 556-812 <SCH>  
R/Malinowski, D.P.; Sadler, J.E.; Davie, E.W.  
Biochemistry 23, 4243-4250, 1984  
A/Title: Characterization of a complementary deoxyribonucleic acid coding for human and  
A/Reference number: I45961; MUID:85023311; PMID:6148961  
A/Accession: I45961  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 706-743, 'R', 745-812 <VAL>  
A/Cross-references: GB:K02935; NID:G163551; PIDN:AAA30714.1; PID:G163552  
R/Brumsholtz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M.;  
Eur. J. Biochem. 114, 465-470, 1981  
A/Title: Comparison of the primary structure of the N-terminal CNBR fragments of human,  
A/Reference number: S03735; MUID:31212097; PMID:7238497  
A/Accession: S03736  
A/Molecule type: protein  
A/Residues: 27-83 <BRU>  
C/Function:  
A/Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a va  
ns the walls of the graafian follicle; also activates the urokinase-type plasminogen act  
A/Pathway: fibrinolysis  
C/Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology  
C/Keywords: duplication; fibrinolysis; glycoprotein; hydrolase; kidney; kringle; plasmin;  
F;1-26/Domain: signal sequence #status predicted <SIG>

F;8-103/Domain: plasminogen-related protein precursor homology <PLPH>  
F;27-812/Product: plasminogen #status experimental <PRO>  
F;27-103/Domain: activation peptide #status experimental <APT>  
F;104-583,584-812/Product: plasmin #status experimental <MAT>  
F;104-583/Domain: plasmin chain A #status experimental <ACH>  
F;110-188/Domain: kringle homology <KR1>  
F;192-265/Domain: kringle homology <KR2>  
F;282-359/Domain: kringle homology <KR3>  
F;384-461/Domain: kringle homology <KR4>  
F;485-564/Domain: kringle homology <KR5>  
F;584-812/Domain: plasmin chain B #status experimental <BCH>  
F;584-805/Domain: trypsin homology <TRY>  
F;56-80,60-68,110-186,131-171,159-183,192-265,195-323,213-252,241-264,282-359,  
3-342  
Done: #status predicted  
F;315/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F;365/Binding site: carbohydrate (Ser) (covalent) #status experimental  
F;624,667,762/Active site: His, Asp, Ser #status predicted

Query Match 21.5%; Score 484.5; DB 1; Length 812;  
Best Local Similarity 35.4%; Pred. No. 9.2e-30;  
Matches 123; Conservative 42; Mismatches 151; Indels 31; Gaps 11  
QY 50 CYEGNHFYRGKASTDTMGPRCLPWNATVLQOYTHAHRSADALQGLGKNNYCRNPD-NR 106  
Db 485 CMIGTKSYRGKATTVAGVPCQEWAAQEPHQHSIFTPETNP-QSGL-ERNYCRNPDGDV 542  
QY 109 RRPWCYVQVGLKPLVQECMHDCADGKLKFCQCKOTLRPR---FKIIGGEFTTIENOPWF 165  
Db 543 NGPWCYTNPRKPF-DYCDVPOC---ESSFDCGPKVPEPKKCSGRIVGGCVSKPHSWPQ 598  
QY 166 AAIYRRHRGGSVTVVCGSLSPCWISATHCFIDYPKEDYIYVLRSLNSNTQGBMK 223  
Db 599 VSLRRSSR---HFCGGTLSPKXWLTAAHCLNIALSFKVILGAHNEKVRQSVC 652  
QY 226 FEVENLILHKDYSADTLAHNDIALKIRSKGRCAPSRITQICLPSMYNDPOFGTSC 281  
Db 654 IPVSLRFPSPQA-----DIALKL-----SRPATITKEVIPACLPFPYVVAARTEC 703  
QY 286 BITGFGKENSTDYLPY-QLAQTWTVVVKLISHRECQPHYGVGSEVTTKMLCAADPQWKTDCS 344  
Db 702 YITGGETQGT---FGEGLLKEAHLPVLENKVRVEYLNVRKSTELCAGHLGGTDCS 758  
QY 345 QDSCGGLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWI 391  
Db 759 QDSCGGLVCFEKDKYILQGVTSWGLGCRPNKPGVYVRSRFDWI 805

## RESULT 28

## S00657

apolipoprotein(a) (EC 3.4.21.-) precursor [validated] - human  
N/Alternate names: apolipoprotein(a); lipoprotein(a) chain apo(a)  
C/Species: Homo sapiens (man)  
C/Date: 30-Jun-1989 #sequence revision 30-Jun-1989 #text\_change 08-Dec-2000  
C/Accession: S00657; A28017; A47277; I60906; A47233; I52415; I65286  
R/McLean, J.W.; Tomlinson, J.E.; Kuang, W.J.; Eaton, D.L.; Chen, E.Y.; Fless, M.; S  
Nature 330, 132-137, 1987  
A/Title: cDNA sequence of human apolipoprotein(a) is homologous to plasminogen  
A/Reference number: S00657; MUID:88039109; PMID:3670400  
A/Accession: S00657  
A/Molecule type: mRNA  
A/Residues: 1-4548 <MCL>  
A/Cross-references: GB:X06290; EMBL:X06696; NID:G28619; PIDN:CAA29618.1; PID:G  
R/Eaton, D.L.; Fless, G.M.; Kohr, W.J.; McLean, J.W.; Xu, Q.T.; Miller, C.G.;  
Proc. Natl. Acad. Sci. U.S.A. 84, 3224-3228, 1987  
A/Title: Partial amino acid sequence of apolipoprotein(a) shows that it is hom  
A/Reference number: A28017; MUID:87204109; PMID:3472206  
A/Accession: A28017  
A/Molecule type: protein  
A/Residues: 20-21, 'P', 23-34, 177-179, 'N', 181-186, 'T', 188-196, 'DKG', 200-292-314,  
X', 4396-4401 <RNT>  
R/Wade, D.P.; Clarke, J.G.; Lindahl, G.E.; Liu, A.C.; Zysow, B.R.; Meer, K.; S  
Proc. Natl. Acad. Sci. U.S.A. 90, 1369-1373, 1993  
A/Title: 5' control regions of the apolipoprotein(a) gene and members of the r